

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:05:20 ; Search time 44.37 Seconds
(without alignments)
20.495 Million cell updates/sec

Title: US-09-276-868-3

Perfect score: 90
Sequence: 1 RRRRRPPYLRPPRP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.0601.*
1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	15	21	AA26885
2	90	100.0	19	17	AAW01452
3	90	100.0	26	17	AAW01447
4	90	100.0	26	19	AAW75723
5	90	100.0	39	14	AA30491
6	90	100.0	39	17	AA30491
7	90	100.0	39	17	AA30491
8	90	100.0	39	17	AA30491
9	90	100.0	39	19	AAW75722
10	90	100.0	39	21	AA26888
11	90	100.0	44	22	AA51194

12	83	92.2	14	17	AAW01450
13	83	92.2	14	19	AAW75725
14	75	83.3	23	17	AAW01451
15	66	73.3	18	16	AAW79211
16	66	73.3	20	16	AAW75730
17	66	73.3	23	16	AAW79209
18	66	73.3	35	16	AAW79212
19	66	73.3	35	19	AAW64400
20	66	73.3	59	21	AAW91699
21	66	73.3	62	22	AAW51197
22	64	71.1	11	21	AAW26886
23	61	67.8	336	17	AAW05520
24	59	65.6	39	21	AAW44779
25	59	65.6	59	17	AAW94448
26	55.5	61.7	497	20	AAW04972
27	54	60.0	230	21	AAW58809
28	54	60.0	955	21	AAW4231
29	53	58.9	18	19	AAW66395
30	53	58.9	18	19	AAW66395
31	53	58.9	18	20	AAW87604
32	53	58.9	18	21	AAW91694
33	53	58.9	18	21	AAW91695
34	53	58.9	18	21	AAW91769
35	53	58.9	18	21	AAW44319
36	53	58.9	19	10	AAW92157
37	53	58.9	174	21	AAW41684
38	53	58.9	428	21	AAW29605
39	52	57.8	18	19	AAW66397
40	52	57.8	18	21	AAW91696
41	52	57.8	354	22	AAW60349
42	52	57.8	354	22	AAW60375
43	52	57.8	439	13	AAW28150
44	52	57.8	464	21	AAW41255
45	52	57.8	478	21	AAW41254

ALIGNMENTS

RESULT 1
ID AAB26885 standard; peptide; 15 AA.
AC AAB26885;
DT 01-FEB-2001 (first entry)
DE PR-39 derived angiogenesis regulatory peptide 1.
KW Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
KW myocardial ischemia; proteasome.
OS Synthetic.
PN WO200057895-A1.
PD 05-OCT-2000.
PF 16-MAR-2000; 2000MO-US07050.
PR 26-MAR-1999; 99US-0276868.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX Simons M, Gao Y;
XX WPI; 2000-628319/60.
XX Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
XX infarction, by administering a PR-39 oligopeptide that regulates
XX enzymatic activity of proteasomes
XX Claim 12; Page 40; 51pp; English.

XX This invention relates to a method for the stimulation of angiogenesis in
CC situ within a targeted collection of viable cells. The method comprises
CC introducing, into the cytoplasm, at least 1 member of the PR-39
CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
CC Part of the proteolytic activity of the proteasomes is selectively
CC altered so as to stimulate angiogenesis. The method is used to induce
CC angiogenesis in tissue that has suffered anoxia or infarction,
CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
CC study the mechanisms that control angiogenesis. The present sequence
CC represents a PR-39 derived peptide which interacts with the proteasome
CC and can be used in the method of the invention.

XX Sequence 15 AA;

Query Match 100.0%; Score 90; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRRPRPPYLPRPP 15
DB 1 rrrprppylprpp 15

RESULT 2

AAW01452
ID AAW01452 standard; peptide; 19 AA.

XX AAW01452;

DE 18-JUN-1997 (first entry)

XX Leukocyte O2- production inhibitor peptide PR19.

XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; p19;
KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
KW tissue damage; oxygen radical; inflammatory disease; therapy.

OS Synthetic.

XX W09632129-A1.

PD 17-OCT-1996.

PF 10-APR-1996; 96WO-US04674.

PR 10-APR-1995; 95US-0419066.

PA (UNIV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Shi J;

DR WPI: 1996-476842/47.

PT Inhibition of leukocyte superoxide anion prodn. and attraction of
XX leukocytes - using peptide(s) partic. based on antimicrobial PR-39
PS Disclosure: Page 27; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
CC isolated from porcine small intestine, and has also been identified in
CC human and porcine neutrophils. PR39 kills bacteria by interfering with
CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
CC mesenchymal cells. Syndecans are important in wound repair, showing that
CC PR39 can be used in wound repair, as well as in antibacterial agents.
CC These sequences, and PR39, can be used in the method of the invention.
CC The method of the invention is for inhibiting leukocyte superoxide anion
CC (O2-) production. The method comprises administering to a leukocyte a
CC peptide (such as this sequence) capable of inhibiting leukocyte O2-

CC production. The peptides can be used as medicaments for fighting
CC infection by attracting leukocytes to a wound site and restricting
CC tissue damage at the wound site caused by excessive oxygen radicals
CC produced by these leukocytes. They can also be used to develop products
CC for treating inflammatory disease states.

XX Sequence 19 AA;

Query Match 100.0%; Score 90; DB 17; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRRPRPPYLPRPP 15
DB 1 rrrprppylprpp 15

RESULT 3

AAW01447
ID AAW01447 standard; peptide; 26 AA.

XX AAW01447;

DE 18-JUN-1997 (first entry)

XX Leukocyte O2- production inhibitor peptide PR26.

XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; p19;
KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
KW tissue damage; oxygen radical; inflammatory disease; therapy.

OS Synthetic.

XX W09632129-A1.

PD 17-OCT-1996.

PF 10-APR-1996; 96WO-US04674.

PR 10-APR-1995; 95US-0419066.

PA (UNIV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Shi J;

DR WPI: 1996-476842/47.

PT Inhibition of leukocyte superoxide anion prodn. and attraction of
XX leukocytes - using peptide(s) partic. based on antimicrobial PR-39
PS Claim 3; Page 26; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
CC isolated from porcine small intestine, and has also been identified in
CC human and porcine neutrophils. PR39 kills bacteria by interfering with
CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
CC mesenchymal cells. Syndecans are important in wound repair, showing that
CC PR39 can be used in wound repair, as well as in antibacterial agents.
CC These sequences, and PR39, can be used in the method of the invention.
CC The method of the invention is for inhibiting leukocyte superoxide anion
CC (O2-) production. The method comprises administering to a leukocyte a
CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
CC production. The peptides can be used as medicaments for fighting
CC infection by attracting leukocytes to a wound site and restricting
CC tissue damage at the wound site caused by excessive oxygen radicals
CC produced by these leukocytes. They can also be used to develop products
CC for treating inflammatory disease states.

XX Sequence 26 AA;

Query Match 100.0%; Score 90; DB 17; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPPYLPRPP 15
 |||||
 Db 1 rrrprpylprpp 15

RESULT 4

AAW5723
 ID AAW5723 standard; peptide; 26 AA.

AC AAW5723;

DT 19-NOV-1998 (first entry)

DE Proline/Arginine rich peptide PR-26.

KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
 KM superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
 KW coronary bypass; organ transplantation surgery.

OS Synthetic.

PN WO9835690-A1.

PD 20-AUG-1998.

PF 17-FEB-1998; 98WO-US03207.

PR 16-FEB-1998; 98US-0024975.
 18-FEB-1997; 97US-0802306.

PA (UNIV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Ross CR, Shi J;

DR WPI; 1998-495359/42.

PT Reduction of reperfusion injury in temporarily occluded blood
 vessels - by administration of a peptide which is rich in proline
 or arginine residues

PS Claim 3; Page 14-15; 35pp; English.

CC Sequences AAW5722-W75732 are proline/arginine rich peptides that upon
 administration into a mammal's bloodstream reduce reperfusion injury
 (production of reactive oxygen species, neutrophil adherence to
 endothelium, and extravasation of neutrophils). These peptides have two
 requirements: they contain the consensus sequence PXXP, where P is a
 proline residue and X is any amino acid residue, which has been found to
 inhibit superoxide production, and secondly they have arginine residues
 adjacent to these motifs, required for effective inhibition. It was
 established by structural and function analysis that a peptide should
 ideally contain 4 or 6 of these motifs, and that inhibitory activity is
 correlated with the increase of length of peptides. The effectiveness
 of these peptides was determined by investigating the production of the
 neutrophil superoxide anion, and also the inhibition of neutrophil
 chemotaxis. From this, it was found that all of the peptides inhibited
 NADPH oxidase to some extent. All of the peptides also inhibit
 neutrophil oxidase activity. PR-39 is believed, to be the most potent
 endogenous down regulator of NADPH oxidase yet discovered, and from the
 data produced, it can be suggested to be involved in eliminating or
 reducing the reperfusion injury induced adhesion and extraction of
 neutrophils. The peptides are also useful in connection with surgical
 procedures such as coronary bypass and organ transplantation surgery.

Sequence 26 AA;
 XX
 XX
 DT

Query Match 100.0%; Score 90; DB 19; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPPYLPRPP 15
 |||||
 Db 1 rrrprpylprpp 15

RESULT 5

AAW0491
 ID AAR30491 standard; peptide; 39 AA.

AC AAR30491;

DT 12-MAY-1993 (first entry)

DE Antibacterial peptide.

KW Pig; small intestine; endocrine; gram negative; bacteria; therapeutic;
 KW veterinary medicine; prophylactic.

OS Sus scrofa domestica.

PN WO9222578-A.

PD 23-DEC-1992.

PF 10-JUN-1992; 92WO-SE00394.

PR 14-JUN-1991; 91SE-0001838.

PA (BOMA/) BOMAN H G.

PA (JOER/) JOERNALL H.

PA (LEE/) LEE J.

PA (MUTT/) MUTT V.

PI Boman HG, Joernvall H, Lee J, Mutt V;

DR WPI; 1993-018080/02.

PT New anti-bacterial polypeptide - active against Gram negative
 bacteria

PS Claim 1; Page 10; 15pp; English.

CC This peptide was isolated from the small intestine of a pig. The
 CC small intestine is an important endocrine organ and many
 CC physiologically active peptides have been isolated from it. This
 CC peptide inhibits the growth of, and may kill, bacteria, pref. gram
 CC negative bacteria. This peptide or its functional derivatives may be
 CC used in human or veterinary medicine for therapeutic or prophylactic
 CC use.

Sequence 39 AA;

Query Match 100.0%; Score 90; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPPYLPRPP 15
 |||||
 Db 1 rrrprpylprpp 15

RESULT 6

AAW01446
 ID AAW01446 standard; peptide; 39 AA.

AC AAW01446;

DT 18-JUN-1997 (first entry)

PD 21-MAR-1996.
 XX
 PF 13-SEP-1995; 95MO-US11675.
 XX
 PR 13-SEP-1994; 94US-0305475.
 XX
 PA (MAGA-) MAGNININ PHARM INC.
 XX
 PI Bedi G, Jacob L, Williams T, Zasloff M;
 XX
 DR WPI; 1996-179725/18.
 XX
 PT Inhibiting sexually transmitted disease e.g. HIV or herpes simplex -
 PT by administering magainin antimicrobial or squalamine cpd. to
 PT inhibit transmission
 XX
 PS Example 1; Page 32; 60pp; English.
 XX
 CC AAR9116-R99123 are antimicrobial, magainin-analogue peptides that may
 CC be used to treat sexually transmitted diseases (STDs) caused by
 CC Chlamydia, HIV, herpes simplex virus, Neisseria gonorrhoeae or
 CC Candida infection. The peptides inhibit STDs by either killing the
 CC infectious organism, impeding the infection mechanism or
 CC interrupting the replication cycle of the organism. Squalamine (an
 CC antinosterol host defense molecule of the dog fish shark Squalus
 CC acanthias) and Pgla (a frog antimicrobial peptide) analogues may
 CC also be useful in inhibiting STD infection and transmission.
 XX
 SQ Sequence 39 AA;
 XX
 Query Match 100.0%; Score 90; DB 17; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPPYLRPRPP 15
 Db 1 rrrprppyldrprpp 15
 XX
 RESULT 9
 AAM75722
 ID AAM75722 standard; peptide; 39 AA.
 XX
 AC AAM75722;
 XX
 DT 19-NOV-1998 (first entry)
 XX
 DE Proline/Arginine rich peptide PR-39.
 XX
 KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
 KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
 KW coronary bypass; organ transplantation surgery.
 XX
 OS Synthetic.
 XX
 PN MO9835690-A1.
 XX
 PD 20-AUG-1998.
 XX
 PF 17-FEB-1998; 98MO-US03207.
 XX
 PR 16-FEB-1998; 98US-0024975.
 PR 18-FEB-1997; 97US-0802306.
 XX
 PA (UNIV) UNIV KANSAS STATE RES FOUND.
 XX
 PI Blecha F, Ross CR, Shi J;
 XX
 DR WPI; 1998-495359/42.
 XX
 PT Reduction of reperfusion injury in temporarily occluded blood
 PT vessels - by administration of a peptide which is rich in proline

PT or arginine residues
 XX
 PS Claim 3; Page 14; 35pp; English.
 XX
 CC Sequences AAM75722-W75722 are proline/arginine rich peptides that upon
 CC administration into a mammal's bloodstream reduce reperfusion injury
 CC (production of reactive oxygen species, neutrophil adherence to
 CC endothelium, and extravasation of neutrophils). These peptides have two
 CC requirements: they contain the consensus sequence PXXP, where P is a
 CC proline residue and X is any amino acid residue, which has been found to
 CC inhibit superoxide production, and secondly they have arginine residues
 CC adjacent to these motifs, required for effective inhibition. It was
 CC established by structural and function analysis that a peptide should
 CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
 CC correlated with the increase of length of peptides. The effectiveness
 CC of these peptides was determined by investigating the production of the
 CC neutrophil superoxide anion, and also the inhibition of neutrophil
 CC chemotaxis. From this, it was found that all of the peptides inhibited
 CC NADPH oxidase to some extent. All of the peptides also inhibit
 CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
 CC endogenous down regulator of NADPH oxidase yet discovered, and from the
 CC data produced, it can be suggested to be involved in eliminating or
 CC reducing the reperfusion injury induced adhesion and extraction of
 CC neutrophils. The peptides are also useful in connection with surgical
 CC procedures such as coronary bypass and organ transplantation surgery.
 XX
 SQ Sequence 39 AA;
 XX
 Query Match 100.0%; Score 90; DB 19; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPPYLRPRPP 15
 Db 1 rrrprppyldrprpp 15
 XX
 RESULT 10
 AAB26888
 ID AAB26888 standard; peptide; 39 AA.
 XX
 AC AAB26888;
 XX
 DT 01-FEB-2001 (first entry)
 XX
 DE PR-39 peptide used in angiogenesis control.
 XX
 DE Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
 KW myocardial ischaemia; proteasome.
 XX
 OS Synthetic.
 XX
 PN MO200057895-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 16-MAR-2000; 200MO-US07050.
 XX
 PR 26-MAR-1999; 99US-0276868.
 XX
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 XX
 PI Simons M, Gao Y;
 XX
 DR WPI; 2000-628319/60.
 XX
 PT Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PT infarction, by administering a PR-39 oligopeptide that regulates
 PT enzymatic activity of proteasomes -
 XX
 PS Disclosure; Page 21; 51pp; English.
 XX

CC This invention relates to a method for the stimulation of angiogenesis in
CC situ within a targeted collection of viable cells. The method comprises
CC introducing, into the cytoplasm, at least 1 member of the PR-39
CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
CC Part of the proteolytic activity of the proteasomes is selectively
CC altered so as to stimulate angiogenesis. The method is used to induce
CC angiogenesis in tissue that has suffered anoxia or infarction,
CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
CC study the mechanisms that control angiogenesis. The present sequence
CC represents the PR-39 peptide from which peptide used in the method of
CC the invention are derived.

XX Sequence 39 AA:

Query Match 100.0%; Score 90; DB 21; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPYLP RPP 15
Db 1 rrrprpylprpp 15

RESULT 11

AAB51194 standard; peptide: 44 AA.

AC AAB51194;

DT 22-MAR-2001 (first entry)

DE E. coli AMP gene PR39 amino acid sequence.

XX Escherichia coli; E. coli; AMP gene; anti-microbial peptide;

KW screening; preservation; food; feed; paint formulation; detergent;

KM cosmetic; medical device; prosthetic implant; disinfectant;

XX microbial infection; tumour.

OS Escherichia coli.

XX WO200073433-A1.

PD 07-DEC-2000.

PF 29-MAY-2000; 2000WO-DK00287.

PR 31-MAY-1999; 99DK-0000766.

XX (NOVO) NOVO NORDISK AS.

PI Kristensen H;

DR WPI; 2001-070965/08.

XX Example 1; Page 34; 59pp; English.

CC The present invention describes a plasmid which is ligated with a pool
CC of nucleotide sequences (NT) linked to an inducible promoter, to express
CC a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
CC optionally linked to a signal (P)), transformed into host cells and
CC cultured in presence of an inducer to induce expression of the NT. A
CC method of screening (1) a pool of nucleotide sequences to select a
CC nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
CC with the pool of NT; (b) transforming host cells which are sensitive to
CC cells to select viable cells; (c) screening the transformed
CC presence of an inducer to induce expression of NT; (e) selecting cells

CC according to the effect of the inducer on cell growth; and (f) recovering
CC NT encoding the peptide from the selected cells. (1) is useful for
CC screening a pool of nucleotide sequences to select a nucleotide sequence
CC encoding a peptide which is an antimicrobial peptide or an antimicrobial
CC enzyme active on bacteria and for finding and preparing a composition for
CC treatment of human or animal. The antimicrobial peptide obtained using
CC (1) may be employed in preservation of e.g. food/feed, paint
CC formulations, detergents, cosmetics, medical devices such as prosthetic
CC implants and also to disinfect and/or kill microbial cells on an object
CC e.g. as an disinfectant for the treatment of biofilm. The peptides are
CC useful for treating microbial infections and/or tumours. Peptides with
CC improved bio-activity can be developed using (1). The peptides have no
CC negative effect on normal mammalian and/or eukaryotic cells. The present
CC sequence represents an AMP (antimicrobial peptide) gene amino acid
CC sequence, which is used in an example from the present invention.

XX Sequence 44 AA:

Query Match 100.0%; Score 90; DB 22; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPYLP RPP 15
Db 3 rrrprpylprpp 17

RESULT 12

AAW01450 standard; peptide: 14 AA.

AC AAW01450;

DT 18-JUN-1997 (first entry)

DE Leukocyte O2- production inhibitor peptide PR14.

XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;

KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;

KM DNA synthesis; protein synthesis; inhibitor; syndecan expression;

XX mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;

KM tissue damage; oxygen radical; inflammatory disease; therapy.

OS Synthetic.

XX WO9632129-A1.

PD 17-OCT-1996.

PF 10-APR-1996; 96WO-US04674.

PR 10-APR-1995; 95US-0419066.

XX (UNIV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Shi J;

DR WPI; 1996-476842/47.

XX Claim 12; Page 28; 45pp; English.

CC AA001447-W01454 represent fragments of the proline-arginine rich
CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
CC isolated from porcine small intestine, and has also been identified in
CC human and porcine neutrophils. PR39 kills bacteria by interfering with
CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
CC mesenchymal cells. Syndecans are important in wound repair, showing that
CC PR39 can be used in wound repair, as well as in antibacterial agents.
CC These sequences, and PR39, can be used in the method of the invention.

CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O₂⁻) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O₂-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.

XX SQ , Sequence 14 AA;

Query Match 92.2%; Score 83; DB 17; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RRP 14
 Db 1 rrrprppylprpr 14

RESULT 13
 AAW75725
 ID AAW75725 standard; peptide; 14 AA.

XX AC AAW75725;

XX DT 19-NOV-1998 (first entry)

XX DE Proline/Arginine rich peptide.

XX KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
 KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
 KW coronary bypass; organ transplantation surgery.

XX OS Synthetic.

XX PN WO9835690-A1.

XX PD 20-AUG-1998.

XX PF 17-FEB-1998; 98WO-US03207.

XX PR 16-FEB-1998; 98US-0024975.

XX PR 18-FEB-1997; 97US-0802306.

XX PA (UNIV) UNIV KANSAS STATE RES FOUND.

XX PI Blecha F, Ross CR, Shi J;

XX DR WPI; 1998-495359/42.

XX PT Reduction of reperfusion injury in temporarily occluded blood
 PT vessels - by administration of a peptide which is rich in proline
 PT or arginine residues

PS Claim 3; Page 15; 35pp; English.

XX Sequences AAW75722-W75732 are proline/arginine rich peptides that upon
 CC administration into a mammal's bloodstream reduce reperfusion injury
 CC (production of reactive oxygen species, neutrophil adherence to
 CC endothelium, and extravasation of neutrophils). These peptides have two
 CC requirements: they contain the consensus sequence PXXP, where P is a
 CC proline residue and X is any amino acid residue, which has been found to
 CC inhibit superoxide production, and secondly they have arginine residues
 CC adjacent to these motifs, required for effective inhibition. It was
 CC established by structural and function analysis that a peptide should
 CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
 CC correlated with the increase of length of peptides. The effectiveness
 CC of these peptides was determined by investigating the production of the
 CC neutrophil superoxide anion, and also the inhibition of neutrophil
 CC chemotaxis. From this, it was found that all of the peptides inhibited
 CC NADPH oxidase to some extent. All of the peptides also inhibit

CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
 CC endogenous down regulator of NADPH oxidase yet discovered, and from the
 CC data produced, it can be suggested to be involved in eliminating or
 CC reducing the reperfusion injury induced adhesion and extraction of
 CC neutrophils. The peptides are also useful in connection with surgical
 CC procedures such as coronary bypass and organ transplantation surgery.

XX SQ Sequence 14 AA;

Query Match 92.2%; Score 83; DB 19; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RRP 14
 Db 1 rrrprppylprpr 14

RESULT 14
 AAW01451
 ID AAW01451 standard; peptide; 23 AA.

XX AC AAW01451;

XX DT 18-JUN-1997 (first entry)

XX DE Leukocyte O₂- production inhibitor peptide PR23.

XX KW Inhibitor; leukocyte O₂- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.

XX OS Synthetic.

XX PN WO9632129-A1.

XX PD 17-OCT-1996.

XX PF 10-APR-1996; 96WO-US04674.

XX PR 10-APR-1995; 95US-0419066.

XX PA (UNIV) UNIV KANSAS STATE RES FOUND.

XX PI Blecha F, Shi J;

XX DR WPI; 1996-476842/47.

XX PT Inhibition of leukocyte superoxide anion prodn. and attraction of
 PT leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 XX Disclosure; Page 27; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
 CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
 CC isolated from porcine small intestine, and has also been identified in
 CC human and porcine neutrophils. PR39 kills bacteria by interfering with
 CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
 CC mesenchymal cells. Syndecans are important in wound repair, showing that
 CC PR39 can be used in wound repair, as well as in antibacterial agents.
 CC These sequences, and PR39, can be used in the method of the invention.
 CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O₂⁻) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O₂-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.

SQ Sequence 23 AA;
Query Match 83.3%; Score 75; DB 17; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 PRPPYLPRPRP 15
Db 1 prppylprprp 12
RESULT 15
AAR79211
ID AAR79211 standard; Peptide; 18 AA.
XX AC AAR79211;
XX DT 05-MAR-1996 (first entry)
XX DE Bactenecin peptide deriv. C, with antibacterial activity.
XX KW Bac5; Bac7; bactenecin; antibacterial; food preservative.
XX OS Synthetic.
XX PN EP665239-A1.
XX PD 02-AUG-1995.
XX PF 26-JAN-1994; 94EP-0200176.
XX PR 26-JAN-1994; 94EP-0200176.
XX PA (CNDR) CONSIGLIO NAZ DELLE RICERCHE.
XX PA (UYTR-) UNIV TRIESTE.
XX PI Gennaro R, Romeo D, Scocchi M, Skerlavaj B;
XX DR WPI; 1995-264826/35.
XX PT New peptide derivs. of bactenecin - useful as antibacterial agents,
XX PT with an activity comparable to Bac5 and Bac7
XX PS Claim 4; Page 5; 7pp; English.
XX CC AAR79209-12 have antibacterial activity which is at least comparable to
XX CC that of the known bactenecins, Bac5 and Bac7, in molar terms. They may
XX CC be used as food preservatives. The peptides may be prepd. using Fmoc
XX CC techniques on an automated continuous flow synthesiser.
XX SQ Sequence 18 AA;
Query Match 73.3%; Score 66; DB 16; Length 18;
Best Local Similarity 85.7%; Pred. No. 0.063;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 RRRPRPPLPRPRP 14
Db 2 rirprprlprprp 15
RESULT 16
AAR75730
ID AAR75730 standard; peptide; 20 AA.
XX AC AAR75730;
XX DT 19-NOV-1998 (first entry)
XX DE Proline/Arginine rich peptide Bac7.
XX KW

KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
KW coronary bypass; organ transplantation surgery.
XX Synthetic.
XX OS WO9835690-A1.
XX PN 20-AUG-1998.
XX PD 17-FEB-1998; 98WO-US03207.
XX PF 16-FEB-1998; 98US-0024975.
XX PR 18-FEB-1997; 97US-0802306.
XX PA (UNIV) UNIV KANSAS STATE RES FOUND.
XX PI Blecha F, Ross CR, Shi J;
XX DR WPI; 1998-495359/42.
XX PT Reduction of reperfusion injury in temporarily occluded blood
XX PT vessels - by administration of a peptide which is rich in proline
XX PT or arginine residues
XX PS Claim 3; Page 18; 35pp; English.
XX CC Sequences AAW75722-W75732 are proline/arginine rich peptides that upon
XX CC administration into a mammal's bloodstream reduce reperfusion injury
XX CC (production of reactive oxygen species, neutrophil adherence to
XX CC endothelium, and extravasation of neutrophils). These peptides have two
XX CC requirements: they contain the consensus sequence PXXP, where P is a
XX CC proline residue and X is any amino acid residue, which has been found to
XX CC inhibit superoxide production, and secondly they have arginine residues
XX CC adjacent to these motifs, required for effective inhibition. It was
XX CC established by structural and function analysis that a peptide should
XX CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
XX CC correlated with the increase of length of peptides. The effectiveness
XX CC of these peptides was determined by investigating the production of the
XX CC neutrophil superoxide anion, and also the inhibition of neutrophil
XX CC chemotaxis. From this, it was found that all of the peptides inhibited
XX CC NADPH oxidase to some extent. All of the peptides also inhibit
XX CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
XX CC endogenous down regulator of NADPH oxidase yet discovered, and from the
XX CC data produced, it can be suggested to be involved in eliminating or
XX CC reducing the reperfusion injury induced adhesion and extraction of
XX CC neutrophils. The peptides are also useful in connection with surgical
XX CC procedures such as coronary bypass and organ transplantation surgery.
XX SQ Sequence 20 AA;
Query Match 73.3%; Score 66; DB 19; Length 20;
Best Local Similarity 85.7%; Pred. No. 0.069;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 RRRPRPPLPRPRP 14
Db 2 rirprprlprprp 15
RESULT 17
AAR79209
ID AAR79209 standard; Peptide; 23 AA.
XX AC AAR79209;
XX DT 05-MAR-1996 (first entry)
XX DE Bactenecin peptide deriv. A, with antibacterial activity.
XX KW Bac5; Bac7; bactenecin; antibacterial; food preservative.
XX XX

OS Synthetic.
 PN EP665239-A1.
 XX 02-AUG-1995.
 PD 26-JAN-1994; 94EP-0200176.
 XX 26-JAN-1994; 94EP-0200176.
 PR (CNRD) CONSIGLIO NAZ DELLE RICERCHE.
 XX (UYTR-) UNIV TRIESTE.
 PA Gennaro R, Romeo D, Scocchi M, Skerlavaj B;
 PI WPI; 1995-264826/35.
 XX New peptide derivs. of bactenecin - useful as antibacterial agents,
 PT with an activity comparable to Bac5 and Bac7
 XX Claim 2; Page 5; 7pp; English.
 PS AAR79209-12 have antibacterial activity which is at least comparable to
 XX that of the known bactenecins, Bac5 and Bac7, in molar terms. They may
 CC be used as food preservatives. The peptides may be prepd. using Fmoc
 CC techniques on an automated continuous flow synthesiser.
 XX Sequence 23 AA;
 SQ

Query Match 73.3%; Score 66; DB 16; Length 23;
 Best Local Similarity 85.7%; Pred. No. 0.078;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RPRP 14
 Db 2 rirprprlprpr 15

RESULT 18
 AAR79212
 ID AAR79212 standard; Peptide; 35 AA.
 XX
 AC AAR79212;
 XX
 DT 05-MAR-1996 (first entry)
 XX
 DE Bactenecin peptide deriv. D, with antibacterial activity.
 XX
 KW Bac5; Bac7; bactenecin; antibacterial; food preservative.
 XX
 OS Synthetic.
 XX
 PN EP665239-A1.
 PD 02-AUG-1995.
 XX
 PF 26-JAN-1994; 94EP-0200176.
 XX
 PR 26-JAN-1994; 94EP-0200176.
 XX
 PA (CNRD) CONSIGLIO NAZ DELLE RICERCHE.
 XX (UYTR-) UNIV TRIESTE.
 PA Gennaro R, Romeo D, Scocchi M, Skerlavaj B;
 PI WPI; 1995-264826/35.
 XX New peptide derivs. of bactenecin - useful as antibacterial agents,
 PT with an activity comparable to Bac5 and Bac7
 XX Claim 5; Page 5; 7pp; English.
 XX

CC AAR79209-12 have antibacterial activity which is at least comparable to
 CC that of the known bactenecins, Bac5 and Bac7, in molar terms. They may
 CC be used as food preservatives. The peptides may be prepd. using Fmoc
 CC techniques on an automated continuous flow synthesiser.
 XX Sequence 35 AA;
 SQ

Query Match 73.3%; Score 66; DB 16; Length 35;
 Best Local Similarity 85.7%; Pred. No. 0.11;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RPRP 14
 Db 2 rirprprlprpr 15

RESULT 19
 AAW66400
 ID AAW66400 standard; peptide; 59 AA.
 XX
 AC AAW66400;
 XX
 DT 12-JAN-1999 (first entry)
 XX
 DE Cationic peptide Bac7.
 XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.
 XX Bos taurus.
 OS
 XX WO9840401-A2.
 PN
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNicol PJ, West MHP;
 XX
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Disclosure; Page 7; 105pp; English.
 XX
 CC AAW66393 to AAW66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 XX
 SQ Sequence 59 AA;

Query Match 73.3%; Score 66; DB 19; Length 59;
 Best Local Similarity 85.7%; Pred. No. 0.18;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RPRP 14
1 ||||| |||||
Db 2 rirprprlprprp 15

RESULT 20

AAAY91699
ID AAY91699 standard; Peptide; 59 AA.

XX AC AAY91699;

XX DT 06-JUN-2000 (first entry)

XX DE Cationic peptide Bac7 amino acid sequence.

XX KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.

XX OS Unidentified.

XX PN WO9965506-A2.

XX PD 23-DEC-1999.

XX PF 14-JUN-1999; 99WO-CA00552.

XX PR 12-JUN-1998; 98US-0096541.

XX PA (MICR-) MICROLOGIX BIOTECH INC.

XX PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX DR WPI; 2000-223549/19.

XX PT Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours

XX PS Disclosure; Page 9; 94pp; English.

XX CC This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC activation of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.

XX SQ Sequence 59 AA;

Query Match 73.3%; Score 66; DB 21; Length 59;

Best Local Similarity 85.7%; Pred. No. 0.18;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RPRP 14
1 ||||| |||||
Db 2 rirprprlprprp 15

RESULT 21

AAAB51197
ID AAB51197 standard; Peptide; 62 AA.

XX AC AAB51197;

XX DT 22-MAR-2001 (first entry)

XX DE

E. coli AMP gene Bac7 amino acid sequence.

XX Escherichia coli; E. coli; AMP gene; anti-microbial peptide;
KW screening; preservation; food; feed; paint formulation; detergent;
KW cosmetic; medical device; prosthetic implant; disinfectant;
KW microbial infection; tumour.

XX OS Escherichia coli.

XX PN WO200073433-A1.

XX PD 07-DEC-2000.

XX PF 29-MAY-2000; 2000WO-DK00287.

XX PR 31-MAY-1999; 99DK-0000766.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Kristensen H;

XX DR WPI; 2001-070965/08.

Screening for nucleotide sequences encoding antimicrobial peptides by
generating peptide libraries in microorganisms, inducing expression of
peptides, selecting cells and recovering nucleotide sequences from
cells

Example 1; Page 34; 59pp; English.

The present invention describes a plasmid which is ligated with a pool
of nucleotide sequences (NT) linked to an inducible promoter, to express
a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
optionally linked to a signal (P)), transformed into host cells and
cultured in presence of an inducer to induce expression of the NT. A
method of screening (I) a pool of nucleotide sequences to select a
nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
with the pool of NT; (b) transforming host cells which are sensitive to
the peptide with the ligated plasmids; (c) screening the viable cells in the
cells to select viable cells; (d) culturing the viable cells in the
presence of an inducer to induce expression of NT; (e) selecting cells
according to the effect of the inducer on cell growth; and (f) recovering
NT encoding the peptide from the selected cells. (I) is useful for
screening a pool of nucleotide sequences to select a nucleotide sequence
encoding a peptide which is an antimicrobial peptide or an antimicrobial
enzyme active on bacteria and for finding and preparing a composition for
treatment of human or animal. The antimicrobial peptide obtained using
(I) may be employed in preservation of e.g. food/feed, paint
formulations, detergents, cosmetics, medical devices such as prosthetic
implants and also to disinfect and/or kill microbial cells on an object
e.g. as an disinfectant for the treatment of biofilm. The peptides are
useful for treating microbial infections and/or tumours. Peptides with
improved bio-activity can be developed using (I). The peptides have no
negative effect on normal mammalian and/or eukaryotic cells. The present
sequence represents an AMP (antimicrobial peptide) gene amino acid
sequence, which is used in an example from the present invention.

XX SQ Sequence 62 AA;

Query Match 73.3%; Score 66; DB 22; Length 62;

Best Local Similarity 85.7%; Pred. No. 0.18;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPPYLP RPRP 14
1 ||||| |||||
Db 4 rirprprlprprp 17

RESULT 22

AAAB26886
ID AAB26886 standard; peptide; 11 AA.

XX AAB26886;
 XX 01-FEB-2001 (first entry)
 XX PR-39 derived angiogenesis regulatory peptide 2.
 DE Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
 KW myocardial ischaemia; proteasome.
 KW Synthetic.
 OS
 XX WO200057895-A1.
 PN
 XX 05-OCT-2000.
 PD
 XX 16-MAR-2000; 2000WO-US07050.
 XX 26-MAR-1999; 99US-0276868.
 XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 PA
 XX Simons M, Gao Y;
 PI
 XX WPI; 2000-628319/60.
 DR
 XX Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PT infarction, by administering a PR-39 oligopeptide that regulates
 PT enzymatic activity of proteasomes -
 XX
 XX Claim 13; Page 41; 51pp; English.
 PS
 XX This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
 CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents a PR-39 derived peptide which interacts with the proteasome
 CC and can be used in the method of the invention.
 XX
 SQ Sequence 11 AA;
 Query Match 71.1%; Score 64; DB 21; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPPYLPR 11
 DB 1 rrrpppylpr 11
 |||||
 RESULT 23
 ID AAW05520 standard; Protein; 336 AA.
 XX
 AC AAW05520;
 XX
 DT 15-JAN-1997 (first entry)
 DE HCMV Toledo strain UL151 protein (clone tol.22).
 XX
 KW CMV; HCMV; vaccine; diagnosis; UL151.
 XX
 OS Human cytomegalovirus Toledo strain.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 47
 FT /note= "unidentified amino acid"

FT Misc-difference 49
 FT /note= "unidentified amino acid"
 FT Misc-difference 115
 FT /note= "unidentified amino acid"
 XX
 PN WO9630387-A1.
 XX
 XX 03-OCT-1996.
 PD
 XX 26-MAR-1996; 96WO-US04100.
 PF
 XX 31-MAR-1995; 95US-0414926.
 PR
 XX (AVIR-) AVIRON.
 PA
 XX Cha T, Spaete R;
 PI
 XX WPI; 1996-455265/45.
 DR N-PSDB; AAT41418.
 DR
 XX New isolated human cytomegalovirus nucleic acid - from Towne and
 PT Toledo strains, used to develop prods. for the diagnosis, prevention
 PT and treatment of human CMV infections
 PT
 XX Claim 5; Page 88-89; 150pp; English.
 PS
 XX Novel protein UL151 (AAW05520) is the product of an open reading
 CC frame found in a novel nucleic acid (AAT41418) isolated from the
 CC Toledo strain of human cytomegalovirus (HCMV). UL151 and other
 CC novel (see also AAW05502-19) and known (see also AAW05500-01) proteins
 CC of the Toledo strain, as well as new proteins (see also AAW05521-24)
 CC from HCMV Towne, can be produced in transformed host cells and used
 CC in the prodn. of subunit vaccines against HCMV. They may be
 CC surface glycoproteins that are immunogenic or responsible for
 CC tissue tropism, or may influence the immune response of an infected
 CC individual.
 XX
 SQ Sequence 336 AA;
 Query Match 57.8%; Score 61; DB 17; Length 336;
 Best Local Similarity 78.6%; Pred. No. 3;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RRP RPYPVLP RP RP 15
 DB 279 rrpipilqrprrp 292
 ||| || |||||
 RESULT 24
 ID AAB44779 standard; Protein; 39 AA.
 XX
 AC AAB44779;
 XX
 DT 12-FEB-2001 (first entry)
 DE
 DE Human secreted protein sequence encoded by gene 18 SEQ ID NO:78.
 XX
 KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
 KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
 KW cerebrovascular disorder; nervous system disorder; ocular disorder;
 KW wound healing; skin aging; food additive; preservative.
 XX
 OS Homo sapiens.
 XX
 XX WO200058336-A1.
 PN
 XX 05-OCT-2000.
 PD
 XX

PF 23-MAR-2000; 2000WO-US07726.
XX
PR 26-MAR-1999; 99US-0126597.
PR 07-JAN-2000; 2000US-0174877.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-602355/57.
DR N-PSDB; AAC79816.
XX
XX Nucleic acid encoding human secreted proteins, used to treat, prevent,
PT ameliorate or diagnose medical conditions such as cancer, and
PT autoimmune diseases -
XX
XX
PS Claim 11; Page 362; 391pp; English.
XX
XX The polynucleotide sequences given in AAC79799 to AAC79848 encode the
CC human secreted proteins given in AAB44762 to AAB44811. AAB44812 to
CC AAB44829 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are used in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities are:
CC immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and polypeptides can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders, cardiovascular
CC disorders, cerebrovascular disorders, angiogenesis, nervous system
CC disorders, infections caused by bacteria, viruses and fungi and ocular
CC disorders. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. AAC79790 to AAC79798 and
CC AAB44761 represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 39 AA;
SQ
Query Match 65.6%; Score 59; DB 21; Length 39;
Best Local Similarity 66.7%; Pred. No. 0.8;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 RRRPRPPYLP RPP 15
Db 17 rscpppytpirp 31
RESULT 25
AAR94448
ID AAR94448 standard; peptide; 59 AA.
XX
AC AAR94448;
XX
XX 05-NOV-1996 (first entry)
XX
XX Synducin peptide (Bac-7) induces syndecan expression.
DE
XX Synducin; induction; expression; syndecan-1; syndecan-4; surface;
KW mesenchymal cell; fibroblast; epithelial; Bac-7; treatment; stasis;
KW decubitus; ulcers; keloids; skin burns; ischemic tissues;
KW hypercoagulation status; prevention; tumour metastasis; restenosis;
KW inhibition; angiogenesis; proliferation; endothelial.
XX
OS Synthetic.

XX Key Location/Qualifiers
FH Peptide 1..13
FT /note= "peptide fragment giving 0 % increase in
FT syndecan-1 expression on NIH 3T3 cells"
FT
FT Peptide 1..18
FT /note= "peptide fragment giving 250 % increase
FT in syndecan-1 expression on NIH 3T3
FT cells"
FT
FT Peptide 1..23
FT /note= "peptide fragment giving 230 % increase
FT in syndecan-1 expression on NIH 3T3
FT cells"
FT
FT Peptide 5..23
FT /note= "peptide fragment giving 10 % increase
FT in syndecan-1 expression on NIH 3T3
FT cells"
FT
FT Peptide 15..42
FT /note= "peptide fragment giving 0% increase in
FT syndecan-1 expression on NIH 3T3 cells"
XX
XX WO9609322-A2.
XX
XX 28-MAR-1996.
XX
XX 22-SEP-1995; 95WO-US12080.
XX
XX 22-SEP-1994; 94US-0310722.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX
XX Bernfield M, Gallo RL;
XX WPI; 1996-188401/19.
XX
XX Modulating mesenchymal interaction by administration of synducin
XX used in the treatment of wounds, tumours, restenosis, etc
XX
XX Claim 4; Page 27; 34pp; English.
XX
XX The present peptide (previously known as an antibacterial peptide),
XX specifically the claimed peptide comprising residues 1-18, is a
XX synducin, which induces the expression of syndecan-1 and syndecan-4
XX on the surface of mesenchymal cells, esp. fibroblasts
XX and epithelial cells. Synducins may be used in the treatment of
XX stasis and decubitus ulcers, keloids, skin burns, ischemic tissues
XX and hypercoagulation states, prevention of tumour metastasis,
XX restenosis inhibition and endothelial cell angiogenesis and
XX proliferation induction.
XX Human microvascular endothelial cells were assayed for syndecan-4
XX expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
XX the present peptide (10 microm) or a blank, to give respective
XX cell surface syndecan-4 values (mOD/m in) of approx. 1.75, 1.70,
XX 2.75 and 0.95.
XX
XX Sequence 59 AA;
SQ

Query Match 65.6%; Score 59; DB 17; Length 59;
Best Local Similarity 84.6%; Pred. No. 1.1;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 RRRPRPPYLP RPP 13
Db 2 rirpriprip 14
RESULT 26
AAY04972
ID AAY04972 standard; Protein; 497 AA.
XX
AC AAY04972;
XX

KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

OS Homo sapiens.

PN WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC76440.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 3164-3166; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 955 AA;

Query Match 60.0%; Score 54; DB 21; Length 955;

Best Local Similarity 60.0%; Pred. No. 49;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 RRRPRPPLPRPP 15

|| ||| :|| ||

Db 262 rrsrpsvpkpspp 276

RESULT 29

AAW66395

ID AAW66395 standard; peptide; 18 AA.

XX AC AAW66395;

XX 12-JAN-1999 (first entry)

XX

DE Cationic peptide apidaecin IA.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;

KW bacterial infection; tolerance; antibacterial; microorganism;

KW bacteria; fungus; parasite; virus.

XX Apis mellifera.

XX WO9840401-A2.

XX 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

XX 25-FEB-1998; 98US-0030619.

XX 10-MAR-1997; 97US-0040649.

XX 20-AUG-1997; 97US-0915314.

XX 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms

XX Disclosure; Page 7; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.

XX Sequence 18 AA;

Query Match 58.9%; Score 53; DB 19; Length 18;

Best Local Similarity 72.7%; Pred. No. 2;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 RPPYLP RP RP 15

|| ||| :|| ||

Db 4 rpvyipqrpp 14

RESULT 30

AAW66396

ID AAW66396 standard; peptide; 18 AA.

XX AC AAW66396;

XX 12-JAN-1999 (first entry)

XX Cationic peptide apidaecin IB.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;

KW bacterial infection; tolerance; antibacterial; microorganism;

KW bacteria; fungus; parasite; virus.

XX Apis mellifera.

XX WO9840401-A2.

XX

PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNicol PJ, West MHP;
 XX
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Disclosure; Page 7; 105pp; English.
 XX
 CC AAW66393 to AAW66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 XX
 SQ Sequence 18 AA;

Query Match 58.9%; Score 53; DB 19; Length 18;
 Best Local Similarity 72.7%; Pred. No. 2;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15
 ||| :|||
 Db 4 rpyipqprpp 14

RESULT 31
 AAW87604
 ID AAW87604 standard; Protein; 18 AA.
 XX
 AC AAW87604;
 XX
 DT 19-MAR-1999 (first entry)
 XX
 DE Antimicrobial peptide Apidaecin I.
 XX
 KW Antimicrobial; fusion; acidic peptide; recombinant; microorganism;
 KW guamerin; basic peptide; Apidaecin I.
 XX
 OS Insecta.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 18
 FT /note= "encoded by ACT"
 XX
 PN WO9854336-A1.
 XX
 PD 03-DEC-1998.
 XX
 XX 28-MAY-1998; 98WO-KR00132.
 XX
 XX 09-APR-1998; 98KR-0013372.
 PR 28-MAY-1997; 97KR-0021312.
 XX
 XX

PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
 PA (SAMY-) SAMYANG GENEX CORP.
 XX
 PI Hong S, Kang MH, Kim JH, Kim S, Lee H, Lee JH;
 XX
 DR WPI; 1999-059844/05.
 DR N-PSDB; AAV83783.
 XX
 PT New method for mass production of antimicrobial peptides - by
 PT constructing fusion genes comprising acidic and antimicrobial
 PT peptide genes and transforming host with vector containing these
 XX
 PS Example 6; Page 17; 52pp; English.
 XX
 CC The invention relates to mass production of antimicrobial peptides. The
 CC method comprises constructing a fusion gene containing a first gene
 CC encoding a negatively charged acidic peptide having at least two cysteine
 CC residues, and a second gene encoding a positively charged basic
 CC antimicrobial peptide. A host microorganism is transformed with a vector
 CC containing the fusion gene and then cultured. The expressed antimicrobial
 CC peptide is then recovered. The method is used to mass produce
 CC antimicrobial peptides in recombinant microorganisms. The inhibitory
 CC effect of the expressed antimicrobial peptide upon the growth of the host
 CC microorganism is considerably reduced by fusing it to the acidic peptide.
 CC Therefore, the use of the fusion gene provides an economic, recombinant
 CC alternative of mass producing antimicrobial peptides, which overcomes the
 CC disadvantages of low-productivity and poor economy, previously
 CC encountered by recombinant and chemical methods. The present sequence
 CC represents the antimicrobial peptide Apidaecin I. The encoding DNA
 CC can be used along with the acidic peptide Guamerin gene in the
 CC construction of the fusion gene.
 XX
 SQ Sequence 18 AA;

Query Match 58.9%; Score 53; DB 20; Length 18;
 Best Local Similarity 72.7%; Pred. No. 2;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15
 ||| :|||
 Db 4 rpyipqprpp 14

RESULT 32
 AAY91694
 ID AAY91694 standard; Peptide; 18 AA.
 XX
 AC AAY91694;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Cationic peptide Apidaecin IA amino acid sequence.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Unidentified.
 XX
 PN WO9965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX

DR WPI; 2000-223549/19.
XX
PT Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
XX
XX
PS Disclosure; Page 8; 94pp; English.
XX
XX This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.
XX
SQ Sequence 18 AA;

Query Match 58.9%; Score 53; DB 21; Length 18;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPYLP RP RP 15
DB 4 RPVYIPQ RP 14
|||:|:|:|:|

RESULT 33
AAAY91695
ID AAY91695 standard; Peptide; 18 AA.
XX
AC AAY91695;
XX
XX
DT 06-JUN-2000 (first entry)
XX
DE Amino acid sequence of cationic peptide Apidaecin IA.
XX
KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX
OS Synthetic.
XX
XX WO9965506-A2.
XX
XX 23-DEC-1999.
XX
XX 14-JUN-1999; 99WO-CA00552.
XX
XX 12-JUN-1998; 98US-0096541.
XX
XX (MICR-) MICROLOGIX BIOTECH INC.
XX
XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX WPI; 2000-223549/19.
XX
XX Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT
PT
XX
XX Disclosure; Page 14; 94pp; English.
XX
XX This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.
XX
SQ Sequence 18 AA;

Query Match 58.9%; Score 53; DB 21; Length 18;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPYLP RP RP 15
DB 4 RPVYIPQ RP 14
|||:|:|:|:|

RESULT 33
AAAY91695
ID AAY91695 standard; Peptide; 18 AA.
XX
AC AAY91695;
XX
XX
DT 06-JUN-2000 (first entry)
XX
DE Amino acid sequence of cationic peptide Apidaecin IA.
XX
KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX
OS Synthetic.
XX
XX WO9965506-A2.
XX
XX 23-DEC-1999.
XX
XX 14-JUN-1999; 99WO-CA00552.
XX
XX 12-JUN-1998; 98US-0096541.
XX
XX (MICR-) MICROLOGIX BIOTECH INC.
XX
XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX WPI; 2000-223549/19.
XX
XX Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT
PT
XX
XX Disclosure; Page 8; 94pp; English.
XX
XX This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The

PS Disclosure; Page -: 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologues or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention.
 CC Note: The present sequence is not shown in the specification, but
 CC was obtained from GenBank.

XX Sequence 354 AA;

Query Match 57.8%; Score 52; DB 22; Length 354;

Best Local Similarity 57.1%; Pred. No. 35;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRPRPPYLPRPRPP 15
 |:|:| :||:| |
 Db 21 rqpqphhlpppppp 34

RESULT 42

AAB60375

ID AAB60375 standard; Protein; 354 AA.

AC AAB60375;

XX 24-APR-2001 (first entry)

XX Human atonal homologue 1 (Hath1) protein, SEQ ID NO:58.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.

XX Homo sapiens.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US15410.

XX 01-JUN-1999; 99US-0137060.

XX 19-JAN-2000; 2000US-0176993.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27283.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or
 PT any of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation -

PS Claim 34; Page -: 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologues or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention.
 CC Note: The present sequence is not shown in the specification, but
 CC was obtained from GenBank.

XX Sequence 354 AA;

Query Match 57.8%; Score 52; DB 22; Length 354;

Best Local Similarity 57.1%; Pred. No. 35;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRPRPPYLPRPRPP 15
 |:|:| :||:| |
 Db 21 rqpqphhlpppppp 34

RESULT 43

AAR28150

ID AAR28150 standard; Protein; 439 AA.

AC AAR28150;

XX 17-MAR-1993 (first entry)

XX Sugar beet chitinase 1.

XX SBC-1; fungicide; anti-fungal agent; extensins.

XX Beta vulgaris cv monova.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /label= leader

FT Domain 27..46

FT /label= hevein_domain

FT Domain 47..178

FT /label= proline-rich

FT /note= "possibly involved in anchoring
 chitinase 1 protein to the cell wall
 after modification of the prolines to
 glycosylated hydroxyprolines, as in
 extensins"

FT Domain 179..416

FT /label= functional_domain

FT Region 417..439

FT /note= "probably directs the protein to
 the vacuole"

XX WO9217591-A.

XX 15-OCT-1992.

XX 07-APR-1992; 92WO-DK00108.

```

PR 08-APR-1991; 91DK-0000616.
XX (DANI-) DANISCO AS.
XX Berglund L, Bojsen K, Mikkelsen JD, Nielsen KK;
XX WPI; 1992-366261/44.
XX N-PSDB; AAQ29965.
XX Sugar beet chitinase 4 and corresponding DNA - inhibits
XX germination and growth of chitin-containing fungi and is used to
XX produce genetically transformed plants
XX Claim 5; Page 164-168; 254pp; English.
XX A clone encoding the chitinase 1 gene was isolated from a
XX sugar beet EMBL3 genomic library. The sequence encodes a protein
XX having 439 amino acid residues. Transgenic plants having increased
XX resistance to nematodes and chitin-containing plant pathogens,
XX partic. fungi, can be produced using genetic constructs containing
XX the chitinase 1 gene. The protein itself can be used in fungicidal
XX compositions.
XX Sequence 439 AA;

Query Match 57.8%; Score 52; DB 13; Length 439;
Best Local Similarity 69.2%; Pred. No. 42;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RRPPTPLPRPP 15
Db 50 rptprptprpp 62

RESULT 44
AAG41255
ID AAG41255 standard; Protein; 464 AA.
AC AAG41255;
XX
XX DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51306.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0123788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 14-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 99US-0132048.

30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
21-MAY-1999; 99US-0135629.
24-MAY-1999; 99US-0136021.
25-MAY-1999; 99US-0136392.
27-MAY-1999; 99US-0136782.
28-MAY-1999; 99US-0137222.
01-JUN-1999; 99US-0137528.
03-JUN-1999; 99US-0137502.
04-JUN-1999; 99US-0137724.
07-JUN-1999; 99US-0138094.
08-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140354.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.

```

RESULT 35
AAY44319
ID AAY44319 standard; peptide; 18 AA.
XX
AC AAY44319;
XX
DT 29-FEB-2000 (first entry)
XX
DE Antimicrobial peptide, Apidaecin I.
XX
KW purf gene; glutamine pyrophosphoribosyl pyrophosphate amidotransferase;
KW purf derivative; fusion partner; antimicrobial peptide; Apidaecin I;
KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct;
KW neutralise; toxicity; pharmaceutical industry; food industry; insect.
XX
OS Apis mellifera.
XX
FH Key Location/Qualifiers
FT Misc-difference 18
FT /note- "Corresponds to ACT codon"
XX
PN WO9964611-A1.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-KR00282.
XX
PR 09-JUN-1998; 98KR-0022117.
PR 14-MAY-1999; 99KR-0017920.
XX
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;
XX
DR WPI; 2000-097542/08.
DR N-PSDB; AAZ29359.
XX
PT New DNA constructs useful for mass production of antimicrobial peptides
PT in microorganism hosts
XX
PS Claim 1; Fig 1; 67pp; English.
XX
CC The present amino acid sequence is an antimicrobial peptide, Apidaecin I
CC derived from the insect Apis mellifera. It is used along with a
CC derivative of purf gene sequence that functions as a fusion partner.
CC A DNA construct that comprises, this antimicrobial peptide encoding
CC sequence and the entire, partial or derivative of purf gene, is used for
CC mass production of the antimicrobial peptide in microorganisms without
CC killing the host cells. Use of the purf gene derivative sequence,
CC neutralises the toxicity of the antimicrobial peptides against the host
CC microorganism. The antimicrobial peptides are useful commercially in the
CC pharmaceutical and food industries.
XX
SQ Sequence 18 AA;

Query Match 58.9%; Score 53; DB 21; Length 18;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15
Db 4 RPYIPQ RP RP 14

RESULT 36
AAP92157
ID AAP92157 standard; peptide; 19 AA.
XX
AC AAP92157;
XX

DT 26-JUN-1990 (first entry)
XX
DE Apidaecin bacteriostatic/bacteriocidal peptide.
XX
KW Bacteriostasis; bacteriocide; honey bee; haemolymph; apidaecin.
XX
OS Apis mellifera.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label=Val or Ile
FT /label=Val or Ile
FT Misc-difference 18 /label=Ile or Leu
FT /label=Ile or Leu
XX
PN EP299828-A.
XX
PD 18-JAN-1989.
XX
PF 29-JUN-1988; 88EP-0401673.
XX
PR 01-JUL-1987; 87EP-0401530.
XX
PA (PLAN-) PLANT GENETIC SYSTEMS NV.
XX
PI Casteels P, Tempst P, Jacobs F, Vaecck M;
XX
DR WPI; 1989-017312/03.
XX
CC Thermostable peptides having bacteriostatic/ bacteriocidal properties -
CC obtd. by isolation from haemolymph of immune honey bees.
XX
PS Claim 10; page 27; 41pp; English.
XX
CC The peptide is isolated from the haemolymph of immune honey bees. It is
CC thermostable and has bacteriocidal and bacteriostatic properties and is
CC active against both Gram-positive and Gram-negative bacteria. Uses
CC include therapy and prophylaxis in man or animal, plant pathogen control
CC and food/feed control. The peptide is distinct from and free from
CC lysozymes cecropins, dipterocins, magainins, and thermostable lysozyme
CC analogues.
CC See also AAP92158-P92163.
XX
SQ Sequence 19 AA;

Query Match 58.9%; Score 53; DB 10; Length 19;
Best Local Similarity 72.7%; Pred. No. 2.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15
Db 5 RPYIPQ RP RP 15

RESULT 37
AAB41684
ID AAB41684 standard; Protein; 174 AA.
XX
AC AAB41684;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1448 polypeptide sequence SEQ ID NO:2896.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN' CORP.

XX Shinkets RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB; AAC75893.

XX Novel nucleic acids and peptides derived from open reading frame X,
 useful for treating e.g. cancers, proliferative disorders,
 neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 2130-2131; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 which represent the human ORFX open reading frames 1 to 3161. The ORFX
 sequences have activities such as: cytostatic; hepatotropic; vulnery;
 antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 osteoarthritic; anticonvulsant; antithrombotic; immunosuppressant;
 immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 antidiabetic; hypotensive; dermatological; immunosuppressive;
 antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 antithyroid; and antianaemic. The sequences can be used for determining
 the presence of or predisposition to, or preventing or treating
 pathological conditions associated with an ORFX-associated disorder. The
 nucleic acids can be used to express ORFX proteins in gene therapy
 vectors. The proteins and nucleic acids may be used to treat cancers,
 proliferative disorders, neurodegenerative disorders, osteoarthritis,
 graft vs host disease, cardiovascular disease, diabetes mellitus,
 hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 174 AA;

Query Match 58.9%; Score 53; DB 21; Length 174;

Best Local Similarity 68.8%; Pred. No. 14;

Matches 11; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

Oy 2 RRRPPYLP--RRPP 15

Db 90 rrprrpprrpprrpp 105

RESULT 38

AAG29606

ID AAG29606 standard; Protein: 428 AA.

XX AAG29606;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 35255.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 KW Arabidopsis thaliana.
 XX EP1033405-A2.
 XX 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 XX 09-MAR-1999; 99US-0123548.
 XX 23-MAR-1999; 99US-0125788.
 XX 25-MAR-1999; 99US-0126284.
 XX 29-MAR-1999; 99US-0126785.
 XX 01-APR-1999; 99US-0127462.
 XX 06-APR-1999; 99US-0128234.
 XX 08-APR-1999; 99US-0128714.
 XX 16-APR-1999; 99US-0129845.
 XX 19-APR-1999; 99US-0130077.
 XX 21-APR-1999; 99US-0130449.
 XX 23-APR-1999; 99US-0130510.
 XX 23-APR-1999; 99US-0130891.
 XX 28-APR-1999; 99US-0131449.
 XX 30-APR-1999; 99US-0132048.
 XX 30-APR-1999; 99US-0132407.
 XX 04-MAY-1999; 99US-0132484.
 XX 05-MAY-1999; 99US-0132485.
 XX 06-MAY-1999; 99US-0132486.
 XX 06-MAY-1999; 99US-0132487.
 XX 07-MAY-1999; 99US-0132863.
 XX 11-MAY-1999; 99US-0134256.
 XX 14-MAY-1999; 99US-0134218.
 XX 14-MAY-1999; 99US-0134219.
 XX 14-MAY-1999; 99US-0134221.
 XX 14-MAY-1999; 99US-0134370.
 XX 18-MAY-1999; 99US-0134768.
 XX 19-MAY-1999; 99US-0134941.
 XX 20-MAY-1999; 99US-0135124.
 XX 21-MAY-1999; 99US-0135353.
 XX 24-MAY-1999; 99US-0135629.
 XX 25-MAY-1999; 99US-0136021.
 XX 27-MAY-1999; 99US-0136392.
 XX 28-MAY-1999; 99US-0136782.
 XX 01-JUN-1999; 99US-0137222.
 XX 03-JUN-1999; 99US-0137528.
 XX 04-JUN-1999; 99US-0137502.
 XX 07-JUN-1999; 99US-0137724.
 XX 08-JUN-1999; 99US-0138094.
 XX 10-JUN-1999; 99US-0138540.
 XX 10-JUN-1999; 99US-0138847.
 XX 14-JUN-1999; 99US-0139119.
 XX 16-JUN-1999; 99US-0139452.
 XX 16-JUN-1999; 99US-0139453.
 XX 17-JUN-1999; 99US-0139492.
 XX 18-JUN-1999; 99US-0139454.
 XX 18-JUN-1999; 99US-0139455.
 XX 18-JUN-1999; 99US-0139456.
 XX 18-JUN-1999; 99US-0139457.
 XX 18-JUN-1999; 99US-0139458.
 XX 18-JUN-1999; 99US-0139459.
 XX 18-JUN-1999; 99US-0139460.
 XX 18-JUN-1999; 99US-0139461.
 XX 18-JUN-1999; 99US-0139462.
 XX 18-JUN-1999; 99US-0139463.
 XX 18-JUN-1999; 99US-0139750.
 XX 18-JUN-1999; 99US-0139763.

Indolicidin analogue; resistance; cationic peptide; antibiologic;
 bacterial infection; tolerance; antibacterial; microorganism;
 bacteria; fungus; parasite; virus.

OS Apis mellifera.

PN WO9840401-A2.

XX 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

XX 25-FEB-1998; 98US-0030619.

PR 10-MAR-1997; 97US-0040649.

PR 20-AUG-1997; 97US-0915314.

PR 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing
PT activity of antibiotic or overcoming tolerance, acquired resistance
PT or inherent resistance of microorganisms

XX Disclosure; Page 7; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the
CC present invention. The present invention describes compositions and
CC methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has
CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.

XX Sequence 18 AA;

Query Match 57.8%; Score 52; DB 19; Length 18;

Best Local Similarity 72.7%; Pred. No. 2.6;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15

DB 4 RPIYIP QP RP 14

RESULT 40

AAY91696

ID AAY91696 standard; Peptide; 18 AA.

XX AC AAY91696;

XX 06-JUN-2000 (first entry)

XX Cationic peptide Apidaecin II amino acid sequence.

XX Cationic peptide: tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.

XX Unidentified.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours

XX Disclosure; Page 8; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 18 AA;

Query Match 57.8%; Score 52; DB 21; Length 18;

Best Local Similarity 72.7%; Pred. No. 2.6;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLP RP RP 15

DB 4 RPIYIP QP RP 14

RESULT 41

AAB60349

ID AAB60349 standard; Protein; 354 AA.

XX AC AAB60349;

XX 24-APR-2001 (first entry)

XX Human atonal homologue 1 (AFOH1, Hath1) protein, SEQ ID NO:2.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Homo sapiens.

XX WO2000073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US15410.

XX 01-JUN-1999; 99US-0137060.

XX 19-JAN-2000; 2000US-0176993.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27252.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or
PT any of its homologs or orthologs, for the treatment of e.g. deafness,
PT osteoarthritis and abnormal cell proliferation

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139877.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

Mon Sep 24 10:12:21 2001

us-09-276-868-3.rat

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model
Run on: September 24, 2001, 10:05:20 : Search time 26.42 seconds
(without alignments)
11.690 Million cell updates/sec

Title: US-09-276-868-3
Sequence: 1 RRRPRPYLRPRPP 15

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 197339 segs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
post-processing: Minimum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/pdata1/1aa/5A_COMB.pep:*
3: /cgn2_6/pdata1/1aa/5B_COMB.pep:*
4: /cgn2_6/pdata1/1aa/6B_COMB.pep:*
5: /cgn2_6/pdata1/1aa/6C_COMB.pep:*
6: /cgn2_6/pdata1/1aa/6D_COMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	90	100.0	26	2	US-08-419-066-2
2	90	100.0	26	4	US-09-024-975-2
3	90	100.0	39	1	US-08-162-072-1
4	90	100.0	39	1	US-08-310-722-1
5	90	100.0	39	2	US-08-419-066-1
6	90	100.0	39	2	US-08-728-333-1
7	90	100.0	39	4	US-09-024-975-1
8	90	100.0	39	5	PCT-US95-12080-1
9	83	92.2	14	4	US-09-024-975-9
10	66	73.3	20	4	US-08-414-926A-26
11	61	67.8	336	2	US-08-926-922-26
12	61	67.8	336	2	US-09-253-682-26
13	61	67.8	336	2	US-09-253-682-26
14	59	65.6	59	5	PCT-US95-12080-3
15	53	58.9	18	1	US-08-205-938A-23
16	53	58.9	18	1	US-08-205-938A-24
17	53	58.9	18	1	US-09-230-180-20
18	53	58.9	18	5	PCT-US95-02626-23
19	53	58.9	18	5	PCT-US95-02626-25
20	52	57.8	18	1	US-08-205-938A-25
21	51.5	57.2	355	4	US-08-483-533-41
22	51.5	57.2	355	4	PCT-US91-06532-3
23	51.5	57.2	355	5	US-08-205-938A-8
24	51	56.7	16	1	PCT-US95-02626-8
25	51	56.7	16	3	US-09-187-331-5
26	51	56.7	16	3	US-09-187-331-1
27	51	56.7	16	3	US-09-187-331-1

28	50.5	56.1	169	4	US-08-483-533-28	Sequence 28, Appl
29	50	55.6	16	1	US-08-205-938A-7	Sequence 7, Appl
30	50	55.6	16	1	US-08-205-938A-28	Sequence 28, Appl
31	50	55.6	16	5	PCT-US95-02626-7	Sequence 26, Appl
32	50	55.6	16	5	PCT-US95-02626-28	Sequence 27, Appl
33	50	55.6	17	1	US-08-205-938A-27	Sequence 27, Appl
34	50	55.6	17	5	PCT-US95-02626-27	Sequence 26, Appl
35	49	54.4	26	4	US-09-024-975-8	Sequence 8, Appl
36	48	53.3	281	2	US-08-815-190A-2	Sequence 2, Appl
37	48	53.3	281	3	US-08-815-190A-2	Sequence 3, Appl
38	48	53.3	281	4	US-09-290-524-3	Sequence 2, Appl
39	48	53.3	281	5	PCT-US95-00362-2	Sequence 2, Appl
40	48	53.3	281	5	5470718-4	Sequence 25, Appl
41	48	53.3	366	6	5352575-5	Sequence 3, Appl
42	48	53.3	402	2	US-09-091-432-2	Sequence 2, Appl
43	48	53.3	520	2	US-08-769-745-36	Sequence 36, Appl
44	47	52.2	21	2	US-08-769-745-36	Sequence 36, Appl
45	47	52.2	43	4	US-09-024-975-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-419-066-2 Application US/08419066
Sequence 2, Appl
Patent No. 5830993
GENERAL INFORMATION:
APPLICANT: Blecha, Frank
INVENTOR: Shi, Jishu
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: John M. Collins, Hovey, Williams, Timmons &
ADDRESS: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,066
FILING DATE: 530
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 23625
TELEPHONE: (816) 474-9050
TELEFAX: (816) 474-9057
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-419-066-2

Query Match 100.0% Score 90; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.7e-05; Indels 0; Gaps 0,
Matches 15; Conservative 0; Mismatches 0;

OY 1 RRRRPPYLRPPRP 15
Db 1 RRRRPPYLRPPRP 15

RESULT 2

US-09-024-975-2
Sequence 2, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-2

Query Match 100.0%; Score 90; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPPYLRPPRP 15
Db 1 RRRRPPYLRPPRP 15

RESULT 3

US-08-162-052-1
Sequence 1, Application US/08162052
Patent No. 5489575
GENERAL INFORMATION:
APPLICANT: LEE, Jong-Youn
APPLICANT: BOMAN, Hans G
APPLICANT: MUTT, VIKTOR
APPLICANT: JORNVALD, Hans
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND THEIR USE
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,052
FILING DATE: 02-JUN-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9101838-2
FILING DATE: 14-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO.92-22578
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 003300-299
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-162-052-1

Query Match 100.0%; Score 90; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPPYLRPPRP 15
Db 1 RRRRPPYLRPPRP 15

RESULT 4

US-08-310-722-1
Sequence 1, Application US/08310722
Patent No. 5654273
GENERAL INFORMATION:
APPLICANT: Gallo, Richard L.
APPLICANT: Klagsbrun, Michael
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/310,722
FILING DATE: 22-SEP-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284

Mon Sep 24 10:12:21 2001

us-09-276-868-3.ra1

REFERENCE/DOCKET NUMBER: CMCC379
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6535
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Multivall, Hans
AUTHORS: Jorntvall, Hans
TITLE: NO. 5654273el Polypeptides and Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
US-08-310-722-1
Query Match 100.0%; Score 90; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05; Indels 0;
Matches 15; Conservative 0; Mismatches 0;
Ox 1 RRRPPPYLP RRRPP 15
DB 1 RRRPPPYLP RRRPP 15
RESULT 5
US-08-419-066-1 Application US/08419066
Sequence 1, 580993
GENERAL INFORMATION:
PATENT NO. 580993
APPLICANT: Shi, Jishu
SYNTHETIC ANTIMICROBIAL PEPTIDE
TITLE OF INVENTION: 2
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS: Hovey, Williams, Timmons &
ADDRESSEE: Collins
STREET: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,066
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 23625
TELEPHONE: (816) 474-9050
TELEFAX: (816) 474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-419-066-1
Query Match 100.0%; Score 90; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05; Indels 0;
Matches 15; Conservative 0; Mismatches 0;
Ox 1 RRRPPPYLP RRRPP 15
DB 1 RRRPPPYLP RRRPP 15
RESULT 6
US-08-728-333-1 Application US/08728333
Sequence 1, 583897
GENERAL INFORMATION:
PATENT NO. 583897
APPLICANT: Gallo, Richard L.
SYNTHETIC ANTIMICROBIAL PEPTIDE
TITLE OF INVENTION: 1
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS: Patrea L. Padst
ADDRESSEE: 110 Peachtree Street, Suite 2800
STREET: Atlanta
CITY: Georgia
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,333
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Patrea L. Padst
REGISTRATION NUMBER: CMCC379
REFERENCE/DOCKET NUMBER:
TELEPHONE: (404)-815-6558
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Multivall, Hans
AUTHORS: Jorntvall, Hans
TITLE: NO. 5863897el Polypeptides and Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
US-08-728-333-1
Query Match 100.0%; Score 90; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05; Indels 0;
Matches 15; Conservative 0; Mismatches 0;

OY 1 RRRRPPYLP RRP 15
| | | | |
DB 1 RRRRPPYLP RRP 15

RESULT 7
US-09-024-975-1
Sequence 1, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-1

Query Match 100.0%; Score 90; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPPYLP RRP 15
| | | | |
DB 1 RRRRPPYLP RRP 15

RESULT 8
PCT-US95-12080-1
Sequence 1, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia

COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Mutt, Viktor
AUTHORS: Jorvall, Hans
TITLE: Novel Polypeptides And Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
PCT-US95-12080-1

Query Match 100.0%; Score 90; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPPYLP RRP 15
| | | | |
DB 1 RRRRPPYLP RRP 15

RESULT 9
US-09-024-975-4
Sequence 4, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:

NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-4

Query Match 92.28; Score 83; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPYLRPRP 14
| | | | | | | | | | | | | | | |
DB 1 RRRPPYLRPRP 14

RESULT 10

US-09-024-975-9
Sequence 9, Application US/09024975
Patent No. 6133233
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-9

Query Match 73.38; Score 66; DB 4; Length 20;
Best Local Similarity 85.7%; Pred. No. 0.024;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPYLRPRP 14
| | | | | | | | | | | | | | | |
DB 2 RRRPPYLRPRP 15

RESULT 11

US-08-414-926A-26
Sequence 26, Application US/08414926A
Patent No. 5721354
GENERAL INFORMATION:
APPLICANT: Spaete, Richard
APPLICANT: Cha, Tai-An
TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,926A
FILING DATE: March 31, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-011/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
CLONE: tol.22
FEATURE:
NAME/KEY: Protein
LOCATION: 1..336
OTHER INFORMATION: /label- u1.51
US-08-414-926A-26

Query Match 67.8%; Score 61; DB 1; Length 336;
Best Local Similarity 78.6%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRRPPYLRPRP 15
| | | | | | | | | | | | | | | |
DB 279 RRRPPYLRPRP 292

RESULT 12

US-08-926-922-26
Sequence 26, Application US/08926922
Patent No. 5925751
GENERAL INFORMATION:
APPLICANT: Spaete, Richard
APPLICANT: Cha, Tai-An
TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Luann Cserr Attorney at Law
STREET: 750 Arimo Avenue

CITY: Oakland
STATE: CA
COUNTRY: USA
ZIP: 94610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/926,922
FILING DATE: September 10, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR 11A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-834-1448
TELEFAX: 510-839-7810
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
IMMEDIATE SOURCE:
CLONE: tol.22
FEATURE:
NAME/KEY: Protein
LOCATION: 1..336
OTHER INFORMATION: /label= UL151
US-08-926-922-26

Query Match 67.8%; Score 61; DB 2; Length 336;
Best Local Similarity 78.6%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 RRRPPYLRPRRP 15
||| ||| ||| |||
DB 279 RRRPPIILQRRPR 292

RESULT 13
US-09-253-682-26
Sequence 26, Application US/09253682
Patent No. 6040170
GENERAL INFORMATION:
APPLICANT: Spaete, Richard
APPLICANT: Cha, Tai-An
TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Luann Cseerr Attorney at Law
STREET: 750 Arimo Avenue
CITY: Oakland
STATE: CA
COUNTRY: USA
ZIP: 94610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/253,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/926,922
FILING DATE: September 10, 1997
ATTORNEY/AGENT INFORMATION:

NAME: Cseerr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR 11A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-834-1448
TELEFAX: 510-839-7810
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
IMMEDIATE SOURCE:
CLONE: tol.22
FEATURE:
NAME/KEY: Protein
LOCATION: 1..336
OTHER INFORMATION: /label= UL151
US-09-253-682-26

Query Match 67.8%; Score 61; DB 3; Length 336;
Best Local Similarity 78.6%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 RRRPPYLRPRRP 15
||| ||| ||| |||
DB 279 RRRPPIILQRRPR 292

RESULT 14
PCT-US95-12080-3
Sequence 3, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-3

Query Match 65.6%; Score 59; DB 5; Length 59;
Best Local Similarity 84.6%; Pred. No. 0.43;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 RRRPPYLRPR 13
||| ||| ||| |||
DB 2 RRRPPIILQRRPR 14

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
APPLICATION NUMBER: US/08/205,938A
CLASSIFICATION: 51A
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-205-938A-24

Query Match
Best Local Similarity 72.7%; Pred. No. 0.72; 1; Indels 0; Gaps 0;
Matches 8; Conservative

OY 5 RPYLPPRP 15 58.9%; Score 53; DB 1; Length 18;
11111111
DB 4 RPYLPQRP 14 72.7%; Pred. No. 0.72; 1; Indels 0; Gaps 0;
Matches 8; Conservative

RESULT 17
US-09-230-180-20
Sequence 20, Application US/09230180
Patent No. 6183992
GENERAL INFORMATION:
APPLICANT: Kim, Sun-Chang
APPLICANT: Lee, Jae Hyun
APPLICANT: Kang, Min Hyun
APPLICANT: Kim, Jeong Hyun
APPLICANT: Hong, Seung-Suh
APPLICANT: Lee, Hyun-Soo
APPLICANT: Samsung Genex Corporation
APPLICANT: Korea Advanced Institute of Science and Technology
TITLE OF INVENTION: METHOD FOR MASS PRODUCTION OF
FILE REFERENCE: 6181/0F135
CURRENT APPLICATION NUMBER: US/09/230,180
CURRENT FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: PCT/KR98/00132
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: KR 13372/1998
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: KR 21312/1997
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 20
LENGTH: 18
TYPE: PPT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Apidaecin I
US-09-230-180-20

Query Match
Best Local Similarity 72.7%; Pred. No. 0.72; 1; Indels 0; Gaps 0;
Matches 8; Conservative

```

OY 5 RPYILRPPRP 15
1111111111
Db 4 RPYILRPPRP 14

RESULT 18
PCT-US95-02626-23
Sequence 23, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
TITLE OF INVENTION: CASTELS, PETER
TITLE OF INVENTION: APIDACIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS: 43
ADDRESS: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-02626-23

Query Match
Best Local Similarity 58.9%; Score 53; DB 5; Length 18;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 5 RPYILRPPRP 15
1111111111
Db 4 RPYILRPPRP 14

RESULT 19
PCT-US95-02626-24
Sequence 24, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
TITLE OF INVENTION: CASTELS, PETER
TITLE OF INVENTION: APIDACIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS: 43
ADDRESS: COOPER & DUNHAM LLP

STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-02626-24

Query Match
Best Local Similarity 58.9%; Score 53; DB 5; Length 18;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 5 RPYILRPPRP 15
1111111111
Db 4 RPYILRPPRP 14

RESULT 20
US-08-205-938A-25
Sequence 25, Application US/08205938A
Patent No. 546671
GENERAL INFORMATION:
APPLICANT: TEMPEST, PAUL
APPLICANT: CASTELS, PETER
TITLE OF INVENTION: APIDACIN-TYPE PEPTIDE ANTIBIOTICS WITH
TITLE OF INVENTION: IMPROVED ACTIVITIES AND/OR DIFFERENT
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS: 43
ADDRESS: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-205-938A-25

Query Match 57.8%; Score 52; DB 1; Length 18;
Best Local Similarity 72.7%; Pred. No. 0.94;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLRPRPP 15
DB 4 RPIYIPQPRPP 14

RESULT 21

PCT-US95-02626-25
Sequence 25, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
APPLICANT: CASTEELS, PETER
TITLE OF INVENTION: APIRACIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
TITLE OF INVENTION: ACTIVITIES AND/OR DIFFERENT ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-02626-25

Query Match 57.8%; Score 52; DB 5; Length 18;
Best Local Similarity 72.7%; Pred. No. 0.94;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 RPPYLRPRPP 15
DB 4 RPIYIPQPRPP 14

RESULT 22

US-08-483-533-41
Sequence 41, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
TITLE OF INVENTION: Diseases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-41

Query Match 57.2%; Score 51.5; DB 4; Length 355;
Best Local Similarity 73.3%; Pred. No. 17;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2 RPPYLRPRPP 15
DB 162 RPPYLRPRPP 176

RESULT 23

PCT-US91-06532-3
Sequence 3, Application PC/TUS9106532
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
TITLE OF INVENTION: Vaccines and Methods

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza Suite 2100
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06532
FILING DATE: 19910910
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gilder, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27373/8235
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/684-5750
TELEFAX: 312/684-9740
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-06532-3

Query Match Similarity	57.2%	Score	51.5	DB	5	Length	355
Best Local	73.3%	Pred. No.	17				
Matches	11	Conservative	0	Mismatches	3	Indels	1
						Gaps	1

```
QY      2 RRRPPPYLPR-PRPP 15
          | | | | | | | |
Db     162 RPPRRPPRRPPRRPP 176
```

RESULT 24
US-08-205-938A-8
; Sequence 8, Application US/08205938A

GENERAL INFORMATION:
APPLICANT: TEMPST, PAUL
APPLICANT: CASTELS, PETER
TITLE OF INVENTION: APIADACIN-TYPE PEPTIDE ANTIBIOTICS WITH
TITLE OF INVENTION: IMPROVED ACTIVITIES AND/OR DIFFERENT
TITLE OF INVENTION: ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P

```

?      REGISTRATION NUMBER: 28.678
?      REFERENCE/DOCKET NUMBER: 45061
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (212)278-0400
?      TELEFAX: (212) 351-0525
?      INFORMATION FOR SEQ ID NO: 8:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 16 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: peptide
?
US-08-205-938A-8

```

Query Match	56.7%;	Score 51;	DB 1;	Length 16;
Best Local Similarity	72.7%;	Pred. No. 1.1;		
Matches	8;	Conservative	1;	Mismatches
			2;	Indels
				Gaps
				0;

OY	5 RPPYLPRPRPP	15
	:	
Db	2 RPTTVVPAPRPP	12

RESULT 25
PCT-US95-02626-8
; Sequence 8, Application PC/TUS9502626
SEQUENCE INFORMATION.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: Pc-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release 1.1.24
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: PCT/US95/02626
:
: FILING DATE:

```

ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-02626-8

Query Match	56.7%	Score 51	DB 5	Length 16
Best Local Similarity	72.7%	Pred. No. 1.1		
Matches	8	Conservative	1	Mismatches 2
				Indels 0
				Gaps 0

CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-205-938A-7

Query Match 55.6%; Score 50; DB 1; Length 16;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYLP RP RP 15
1 1 1 1 1 1 1 1
DB 2 RPYLP RP RP 12

RESULT 30
US-08-205-938A-28
Sequence 28, Application US/08205938A
Patent No. 5466671
GENERAL INFORMATION:
APPLICANT: TEMPEST, PAUL
TITLE OF INVENTION: APIRACIN-TYPE PEPTIDE ANTIBIOTICS WITH
TITLE OF INVENTION: IMPROVED ACTIVITIES AND/OR DIFFERENT
TITLE OF INVENTION: ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-205-938A-28

Query Match 55.6%; Score 50; DB 1; Length 16;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYLP RP RP 15
1 1 1 1 1 1 1 1
DB 2 RPYLP RP RP 12

RESULT 31
PCT-US95-02626-7
Sequence 7, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
TITLE OF INVENTION: APIRACIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
TITLE OF INVENTION: ACTIVITIES AND/OR DIFFERENT ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-02626-7

Query Match 55.6%; Score 50; DB 5; Length 16;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYLP RP RP 15
1 1 1 1 1 1 1 1
DB 2 RPYLP RP RP 12

RESULT 32
PCT-US95-02626-28
Sequence 28, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
APPLICANT: CASTEELS, PETER
TITLE OF INVENTION: APIDAEICIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
TITLE OF INVENTION: ACTIVITIES AND/OR DIFFERENT ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0525
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-02626-28

Query Match 55.6%; Score 50; DB 5; Length 16;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 RPYLPPRPP 15
1111111111
Db 2 RPYLPPRPP 12

RESULT 33
US-08-205-938A-27
Sequence 27, Application US/08205938A
GENERAL INFORMATION:
APPLICANT: TEMPEST, PAUL
APPLICANT: CASTEELS, PETER
TITLE OF INVENTION: APIDAEICIN-TYPE PEPTIDE ANTIBIOTICS WITH
TITLE OF INVENTION: IMPROVED ACTIVITIES AND/OR DIFFERENT
TITLE OF INVENTION: ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA

ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,938A
FILING DATE: 02-MAR-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0525
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-205-938A-27

Query Match 55.6%; Score 50; DB 1; Length 17;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 RPYLPPRPP 15
1111111111
Db 3 RPYLPPRPP 13

RESULT 34
PCT-US95-02626-27
Sequence 27, Application PC/TUS9502626
GENERAL INFORMATION:
APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
APPLICANT: TEMPEST, PAUL
APPLICANT: CASTEELS, PETER
TITLE OF INVENTION: APIDAEICIN-TYPE PEPTIDE ANTIBIOTICS WITH IMPROVED
TITLE OF INVENTION: ACTIVITIES AND/OR DIFFERENT ANTIBACTERIAL SPECTRUM
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/205,938
FILING DATE: 02-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45061-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0525

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/815,190A
APPLICATION NUMBER: US/08/815,190A
FILING DATE: 11-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/614,584
FILING DATE: 13-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-815-190A-2

Query Match 53.3%; Score 48; DB 3; Length 281;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRRPPPYLPRRPP 15
:|||||
Db 42 QRRPPPPPPPLPP 56

RESULT 38
US-09-290-640-25
Sequence 25, Application US/09290640
Patent No. 6204055
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
APPLICANT: Marcusson, Eric G.
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
FILE REFERENCE: ISPH-0351
CURRENT APPLICATION NUMBER: US/09/290,640
CURRENT FILING DATE: 1999-04-12
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 25
LENGTH: 281
TYPE: PRT
ORGANISM: Homo sapiens
US-09-290-640-25

Query Match 53.3%; Score 48; DB 4; Length 281;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRRPPPYLPRRPP 15
:|||||
Db 42 QRRPPPPPPPLPP 56

RESULT 39
US-09-479-524-3
Sequence 3, Application US/09479524
Patent No. 6268350
GENERAL INFORMATION:
APPLICANT: Barbera-Guillen, Emilio
TITLE OF INVENTION: Polynucleotides For Inhibiting Metastasis And Tumor Cell Growth
FILE REFERENCE: B-37
CURRENT APPLICATION NUMBER: US/09/479,524
CURRENT FILING DATE: 2000-01-07
EARLIER APPLICATION NUMBER: US 09/170,948

EARLIER FILING DATE: 1998-10-13
EARLIER APPLICATION NUMBER: US 60/062,733
EARLIER FILING DATE: 1997-10-23
NUMBER OF SEQ ID NOS: 8
SOFTWARE: WordPerfect
SEQ ID NO 3
LENGTH: 281
TYPE: PRT
ORGANISM: Homo sapiens
US-09-479-524-3

Query Match 53.3%; Score 48; DB 4; Length 281;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRRPPPYLPRRPP 15
:|||||
Db 42 QRRPPPPPPPLPP 56

RESULT 40
PCT-US95-00362-2
Sequence 2, Application PC/TUS9500362
GENERAL INFORMATION:
APPLICANT: IMMUNEX CORPORATION
TITLE OF INVENTION: Ligand That Binds Fas Antigen
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00362
FILING DATE: 06-JAN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,138
FILING DATE: 07-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,559
FILING DATE: 01-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2805-WO
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00362-2

Query Match 53.3%; Score 48; DB 5; Length 281;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRRPPPYLPRRPP 15
:|||||
Db 42 QRRPPPPPPPLPP 56

RESULT 41

5470718-4
; Patent No. 5470718
; APPLICANT: O'CALLAGHAN, DENNIS J.
; TITLE OF INVENTION: EQUINE HERPESVIRUS TYPE 1 GLYCOPROTEIN
; D NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/954,417
; FILING DATE: 30-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 561,553
; FILING DATE: 01-AUG-1990
; SEQ ID NO: 4:
; LENGTH: 366
5470718-4

Query Match 53.3%; Score 48; DB 6; Length 366;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRRPPYLRPRP 14
||| |||
Db 267 RRRPPYLRPRP 278

RESULT 42
5352575-5
; Patent No. 5352575
; APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.
; TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/513,282
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 100,817
; FILING DATE: 29-JUN-1987
; APPLICATION NUMBER: 886,260
; FILING DATE: 16-JUL-1986
; APPLICATION NUMBER: 784,787
; FILING DATE: 04-OCT-1985
; APPLICATION NUMBER: 801,799
; FILING DATE: 26-NOV-1985
; APPLICATION NUMBER: 844,113
; FILING DATE: 26-MAR-1986
; SEQ ID NO: 5:
; LENGTH: 402
5352575-5

Query Match 53.3%; Score 48; DB 6; Length 402;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRRPPYLRPRP 14
||| |||
Db 268 RRRPPYLRPRP 279

RESULT 43
US-09-091-432-2
; Sequence 2, Application US/09091432
; Patent No. 5981837
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
; FILE REFERENCE: 7024-323
; CURRENT APPLICATION NUMBER: US/09/091,432
; CURRENT FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: PCT/US96/20094
; EARLIER FILING DATE: 1996-12-19
; EARLIER APPLICATION NUMBER: US 60/009,119

; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: US 60/013,388
; EARLIER FILING DATE: 1996-03-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Microsoft Word 2.0C
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sequence is deduced from DNA sequence of SEQ ID NO:1
; Patent No. 5981837
US-09-091-432-2

Query Match 53.3%; Score 48; DB 2; Length 520;
Best Local Similarity 75.0%; Pred. No. 62;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRRPPYLRPRP 13
||| |||
Db 34 RRRPPYLRPRP 45

RESULT 44
US-08-769-745-36
; Sequence 36, Application US/08769745
; Patent No. 5955259
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd C.
; APPLICANT: Levitan, Irwin B.
; APPLICANT: Brandeis University
; TITLE OF INVENTION: Mechanism for the Regulation of Ion
; FILE REFERENCE: BRU96-02
; CURRENT APPLICATION NUMBER: US/08/769,745
; CURRENT FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Mouse
US-08-769-745-36

Query Match 52.2%; Score 47; DB 2; Length 21;
Best Local Similarity 61.5%; Pred. No. 4.2;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 RRRPPYLRPRP 15
||| |||
Db 4 KRRPPYLRPRP 16

RESULT 45
US-09-024-975-7
; Sequence 7, Application US/09024975
; Patent No. 6133233
; GENERAL INFORMATION:
; APPLICANT: ROSS, CHRISTOPHER R.
; APPLICANT: BLECHA, FRANK
; APPLICANT: SHI, JISHU
; TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
; STREET: 2405 GRAND BLVD., SUITE 400
; CITY: KANSAS CITY
; STATE: MO
; COUNTRY: USA
; ZIP: 64108
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-7

Query Match 52.2%; Score 47; DB 4; Length 43;
Best Local Similarity 62.5%; Pred. No. 8.1;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
OY 2 RRP--RPYLP RRP 15
||| |||: | |||
Db 7 RRPPIRPFP RRP 22

Search completed: September 24, 2001, 10:06:40
Job time: 80 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:05:20 : Search time 29.48 Seconds
(Without alignments)
38.759 Million cell updates/sec

Title: US-09-276-868-3

Perfect score: 90

Sequence: 1 RRRPRPPLRRPRPP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	90	100.0	172	2 S68232	antimicrobial prot
2	66	73.3	59	2 A36589	bactenecin 7 - boy
3	59.5	66.1	82	2 A41051	spore coat protein
4	58	64.4	190	2 S68230	antimicrobial pept
5	56.5	62.8	168	2 S35330	apidaecin 14 precu
6	56.5	62.8	199	2 S14981	extensin class I (
7	54.5	60.6	301	2 J01663	hybrid proline-ric
8	54	60.0	359	2 T13478	hypothetical prote
9	54	60.0	427	2 T32652	hypothetical prote
10	53	58.9	26	2 S06675	apidaecin 1b precu
11	53	58.9	144	2 S35331	apidaecin 22 precu
12	53	58.9	184	2 T29373	hypothetical prote
13	53	58.9	283	2 S35332	apidaecin 73 precu
14	53	58.9	428	2 E71415	probable coil wall
15	53	58.9	491	2 T07598	proline-rich prote
16	52	57.8	261	1 WMBEXE	infected cell prot
17	52	57.8	439	2 S51939	chitinase (EC 3.2.
18	52	57.8	467	2 S71169	protein kinase, 54
19	52	57.8	1006	2 G86392	hypothetical prote
20	51.5	57.2	1187	1 J04155	protein-tyrosine-p
21	51.5	57.2	1189	1 J04155	protein-tyrosine-p
22	51	56.7	180	2 S43791	PBDX protein - hum
23	50.5	56.1	1216	2 J01005	synaptojanin 2 alp
24	50	55.6	192	2 S76867	hypothetical prote
25	50	55.6	383	2 T06753	zinc finger protei
26	50	55.6	415	1 A34170	acrosin (EC 3.4.21
27	50	55.6	421	2 S29599	acrosin (EC 3.4.21
28	50	55.6	424	2 A54964	spliceosome-associ
29	50	55.6	449	2 S16748	proline-rich prote

30	50	55.6	547	2 C96828	unknown protein F1
31	50	55.6	1460	1 EDBE1F	immediate-early pr
32	50	55.6	3036	2 T18995	hypothetical prote
33	49.5	55.0	589	2 T29299	hypothetical prote
34	49	54.4	118	2 T19345	hypothetical prote
35	49	54.4	161	2 J05572	proline-rich prote
36	49	54.4	210	2 F72593	hypothetical prote
37	49	54.4	210	2 T33700	hypothetical prote
38	49	54.4	218	2 T22261	hypothetical prote
39	49	54.4	256	2 A27319	gliadin - wheat
40	49	54.4	296	2 S07361	alpha/beta-gliadin
41	49	54.4	352	2 F84799	hypothetical prote
42	49	54.4	369	2 S20500	hydroxyproline-ric
43	49	54.4	380	2 T32944	hypothetical prote
44	49	54.4	444	2 E83802	hypothetical prote
45	49	54.4	1203	2 T21275	hypothetical prote

ALIGNMENTS

RESULT 1
S68232
antimicrobial protein PR-39 precursor, cathelin-associated - pig
N:Alternate names: myeloid antibacterial protein PR-39
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C/Accession: S68232; JN0899; I47138; S19563
R/Zhao, C.; Ganz, T.; Lehrer, R.I.
FEBS Lett. 376, 130-134, 1995
A>Title: Structures of genes for two cathelin-associated antimicrobial peptides: prop
A:Reference number: S68232; MUID:96105365
A/Accession: S68232
A/Status: translation not shown
A/Molecule type: DNA
A:Residues: 1-172 <ZHA>
A:Cross-references: EMBL:X69201; NID:g1165150; PIDN:CAA61487.1; PID:g1165151
A:Title: A cDNA derived from pig bone marrow cells predicts a sequence identical to t
A:Reference number: JN0899; MUID:94071853
A/Accession: JN0899
A/Molecule type: mRNA
A:Residues: 1-20, 'A', 22-172 <STO>
A:Cross-references: GB:I23825; NID:g435100; PIDN:AAA31109.1; PID:g435101
A:Experimental source: bone marrow cells
R:Schmidt, G.H.; Magnusson, K.P.; Chowdhary, B.P.; Johansson, M.; Andersson, L.;
Proc. Natl. Acad. Sci. U.S.A. 92, 7085-7089, 1995
A>Title: Structure of the gene for porcine peptide antibiotic PR-39, a cathelin gene
A:Reference number: I47138; MUID:95350216
A/Accession: I47138
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A:Residues: 1-28, 'T', 30-89, 'OR', 92-116, 'NDP', 120-172 <GUD>
A:Cross-references: EMBL:X67236; NID:8629142; PIDN:CAA60682.1; PID:g1051298
R:Agrether, B.; Lee, U.Y.; Bergman, T.; Carlgust, M.; Boman, H.G.; Mutt, V.; Joernv
Eur. J. Biochem. 202, 849-854, 1991
A>Title: Amino acid sequence of PR-39. Isolation from pig intestine of a new member o
A:Reference number: S19563; MUID:92111534
A/Accession: S19563
A/Molecule type: protein
A:Residues: 131-169 <AGE>
A:Experimental source: Intestine
C:Genetics:
A:Gene: PR39
A:Introns: 66/3; 102/3; 126/3
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end; antibacterial
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-129/Domain: cystatin homology <CYS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-169/Product: antimicrobial protein PR-39 #status experimental <MAT>

A:Reference number: S35330; MUID:93223697

A:Accession: S35330

A:Molecule type: mRNA

A:Residues: 1-168 <CAS>

A:Cross-references: EMBL:X75575; NID:g297062; PIDN:CAA51167.1; PID:g297063

R:Castels, P.; Ampe, C.; Jacobs, F.; Vaeck, M.; Tempst, P.

EMBO J. 8, 2387-2391, 1989

A:Title: Apidaecins: antibacterial peptides from honeybees.

A:Reference number: S05383; MUID:90005446

A:Accession: S06676

A:Molecule type: protein

A:Residues: 43-60 <CA2>

C:Superfamily: procylic acid repetitive protein

F:43-60/Product: apidaecin II #status experimental <MAT>

Query Match 62.8%; Score 56.5; DB 2; Length 168;

Best Local Similarity 50.0%; Pred. No. 2.4;

Matches 11; Conservative 2; Mismatches 2; Indels 7; Gaps 1;

QY 1 RRRP-----RPPYLRPRRP 15

DB 117 RREPPEPGNNRPYIPQPRRP 138

RESULT 6

S14981

extensin class I (clone w1-8 L) - tomato (fragment)

C:Species: Lycopersicon esculentum (tomato)

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2000

A:Accession: S14981

R:Shawalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.

Plant Mol. Biol. 16, 547-563, 1991

A:Title: Tomato extensin and extensin-like cDNAs: structure and expression in response t

A:Reference number: S14970; MUID:91329690

A:Accession: S14981

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-199 <SHO>

A:Cross-references: EMBL:X55692

A:Experimental source: cv. UC82B

C:Superfamily: hydroxyproline-rich glycoprotein

C:Keywords: cell wall; glycoprotein; hydroxyproline

Query Match 62.8%; Score 56.5; DB 2; Length 199;

Best Local Similarity 73.3%; Pred. No. 2.8;

Matches 11; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 4 PRPP---YLPRRPP 15

DB 77 PRPPPEYLPRPP 91

RESULT 7

J01663

hybrid proline-rich protein - maize

C:Species: Zea mays (maize)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Sep-1999

A:Accession: J01663

R:Jose-Stanoyl, M.; Ruiz-Avila, L.; Pulgomech, P.

Plant Cell 4, 413-423, 1992

A:Title: A maize embryo-specific gene encodes a proline-rich and hydrophobic protein.

A:Reference number: J01663; MUID:92361259

A:Accession: J01663

A:Molecule type: DNA

A:Residues: 1-301 <JOS>

A:Cross-references: EMBL:X60432; NID:g433706; PIDN:CAA42959.1; PID:g433707

A:Experimental source: strain W64A

C:Superfamily: hydroxyproline-rich glycoprotein

Query Match

60.6%; Score 54.5; DB 2; Length 301;

Best Local Similarity 71.4%; Pred. No. 7.4;

Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 RRRPYL-PRRPP 15

DB 149 RSPPYVPTPRPP 162

RESULT 8

T13478

hypothetical protein 34F3.10 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

A:Accession: T13478

R:Valenti, P.; Salles, C.; Campbell, L.; Glover, D.

submitted to the EMBL Data Library, April 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17685

A:Accession: T13478

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-359 <PHI>

A:Cross-references: EMBL:AL031583; NID:e1321005; PID:e1321018; PIDN:CA841346.1

C:Genetics:

A:Cross-references: FLYBase:FBgn0025623

A:Introns: 17/2; 50/3; 333/2

A>Note: EG:34F3.10

Query Match 60.0%; Score 54; DB 2; Length 359;

Best Local Similarity 71.4%; Pred. No. 10;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRRPYLPRRPP 15

DB 167 RRRPPPLPRPP 180

RESULT 9

T32652

hypothetical protein F39C12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000

A:Accession: T32652

R:Chisoe, S.; Sansone, J.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid F39C12.

A:Reference number: Z21206

A:Accession: T32652

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-427 <CHD>

A:Cross-references: EMBL:AF039043; PIDN:AAB94196.1; GSPDB:GN00028; CESP:F39C12.3

A:Experimental source: strain Bristol N2; clone F39C12

C:Genetics:

A:Gene: CESP:F39C12.3

A:Map position: X

A:Introns: 42/3; 104/3; 133/3; 164/3; 213/3; 276/3; 336/3

Query Match 60.0%; Score 54; DB 2; Length 427;

Best Local Similarity 69.2%; Pred. No. 12;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RRRPYLPRRPP 15

DB 338 RRRPPDIPPLPP 350

RESULT 10

S06675

apidaecin 1b precursor - honeybee

C:Species: Apis mellifera (honeybee)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Dec-1998
C/Accession: S06675

R/Casteels, P.; Ampe, C.; Jacobs, F.; Vaeck, M.; Tempst, P.

EMBO J. 8, 2387-2391, 1989

A>Title: Apidaecins: antibacterial peptides from honeybees.

A/Reference number: S05383; MUID:90005446

A/Accession: S06675

A/Molecule type: protein

A/Residues: 1-26 <CAS>

F:1-8/Domain: propeptide #status experimental <PRO>

F:9-26/Product: apidaecin Ib #status experimental <MAT>

Query Match 58.9%; Score 53; DB 2; Length 26;
Best Local Similarity 72.7%; Pred. No. 1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLRPRRP 15
|||:|||||

DB 12 RPYIIPQPRRP 22

RESULT 11

S35331

apidaecin 22 precursor - honeybee

C/Species: Apis mellifera (honeybee)

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000

C/Accession: S35331

R/Casteels-Josson, K.; Capaci, T.; Casteels, P.; Tempst, P.

EMBO J. 12, 1569-1578, 1993

A>Title: Apidaecin multipetide precursor structure: a putative mechanism for amplifica

A/Reference number: S35330; MUID:93223697

A/Accession: S35331

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-144 <CAS>

A/Cross-references: EMBL:X72576; NID:g297064; PIDN:CAA51168.1; PID:g297065

C/Superfamily: procyelic acidic repetitive protein

Query Match 58.9%; Score 53; DB 2; Length 144;
Best Local Similarity 72.7%; Pred. No. 5.4;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLRPRRP 15
|||:|||||

DB 46 RPYIIPQPRRP 56

RESULT 12

T29373

hypothetical protein ZC404.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C/Accession: T29373

R/Bentley, D.; Le, T.T.

A/Description: The sequence of C. elegans cosmid ZC404.

A/Reference number: Z20614

A/Accession: T29373

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-184 <BEN>

A/Cross-references: EMBL:U55363; PIDN:AAA97967.1; GSPDB:GN00023; CESP:ZC404.1

C/Genetics: C:gene: CESP:ZC404.1

A:Map position: 5

A:introns: 15/2; 50/2; 75/2; 138/2

C:superfamily: Caenorhabditis elegans hypothetical protein ZC404.1

Query Match 58.9%; Score 53; DB 2; Length 184;

Best Local Similarity 90.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPYLRPRRP 12
|||:|||||

DB 26 RPKKPLPRRP 35

RESULT 13

S35332

apidaecin 73 precursor - honeybee (fragment)

N/Contains: apidaecin Ia

C/Species: Apis mellifera (honeybee)

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Nov-2000

C/Accession: S35332; S05383

R/Casteels-Josson, K.; Capaci, T.; Casteels, P.; Tempst, P.

EMBO J. 12, 1569-1578, 1993

A>Title: Apidaecin multipetide precursor structure: a putative mechanism for amplifi

A/Reference number: S35330; MUID:93223697

A/Accession: S35332

A/Molecule type: mRNA

A/Residues: 1-283 <CAS>

A/Cross-references: EMBL:X72577; NID:g297066; PIDN:CAA51169.1; PID:g4539289

A/Accession: S05383

A/Molecule type: protein

A/Residues: 258-283 <CA3>

C/Superfamily: proline-rich protein

F:266-283/Product: apidaecin Ia #status experimental <MAT>

Query Match 58.9%; Score 53; DB 2; Length 283;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RPPYLRPRRP 15
|||:|||||

DB 45 RPYIIPQPRRP 55

RESULT 14

E71415

probable coll wall protein - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

A/Variety: Columbia

C/Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

C/Accession: E71415

R/Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B

Nature 391, 485-488, 1998

A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen

erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.;

C.; Chalwatzis, N.

A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t

A/Reference number: A71400; MUID:98121113

A/Accession: E71415

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-428 <BEV>

A/Cross-references: GB:297338; NID:g2244870; PID:e327461; PID:g2244874

C/Genetics: C:gene: CESP:ZC404.1

A:Map position: 4COP9-4G3845

Query Match 58.9%; Score 53; DB 2; Length 428;
Best Local Similarity 61.5%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPYLRPRRP 15
|||:|||||

DB 67 KPPYIPCPPPP 79

```
RESULT 15
T07598
A:Molecule type: DNA
A:Residues: 1-439 <BER>
A:Cross-references: EMBL:X79301, NID:g488730, PID:g488731
C:Species: Solanum tuberosum (potato)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
C:Accession: T07598
R:Menke, D.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z16040
A:Accession: T07598
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-491 <MEN>
A:Cross-references: EMBL:AJ000997, NID:e1314336, PIDN:CAA04449.1, PID:e1314337
C:Experimental source: cv. Desiree
C:Genetics:
A:Gene: gppl

Query Match          58.8%; Score 53; DB 2; Length 491;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPYLRPRPP 15
DB 306 KRCPSPVPRKRP 319

RESULT 16
WMBEXE
Infected cell protein ICP34.5 - human herpesvirus 2 (strain HGS2)
N:Alternate names: RL1 protein
C:Species: human herpesvirus 2
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C:Accession: J01502
R:McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.
J. Gen. Virol. 72, 3057-3075, 1991
A:Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
A:Reference number: J01502; MUID:92113549
A:Accession: J01502
A:Molecule type: DNA
A:Residues: 1-261 <MCG>
A:Cross-references: GB:J010471, DDBJ:DD01128; NID:g221784; PIDN:BAA23428.1; PID:g2626943
C:Genetics:
A:Gene: RL1
A:Introns: 165/3
C:Superfamily: herpesvirus infected cell protein ICP34.5
C:Keywords: tandem repeat
F:3-12/Region: 5-residue repeats (R-R-P-G-A-P-A-V)
F:16-31/Region: 8-residue repeats (P-R-P-G-A-P-A-V)

Query Match          57.8%; Score 52; DB 1; Length 261;
Best Local Similarity 64.7%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 RRRPR--PYLRPRPP 15
DB 13 RRRPRGAPAVPRGAP 29

RESULT 17
S51939
Chitinase (EC 3.2.1.14) precursor - beet
C:Species: Beta vulgaris (beet)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 10-Jul-1998
C:Accession: S51939; S72315; S45025
R:Berghlund, L.; Brunstedt, J.; Nielsen, K.K.; Chen, Z.; Mikkelsen, J.D.; Marcker, K.A.
Plant Mol. Biol. 27, 211-216, 1995
A:Title: A proline-rich chitinase from Beta vulgaris.
A:Reference number: S51939; MUID:95170004
A:Accession: S51939
```

```
A:Molecule type: DNA
A:Residues: 1-439 <BER>
A:Cross-references: EMBL:X79301, NID:g488730, PID:g488731
A:Note: the authors translated the codon TGC for residue 416 as GLY
A:Accession: S72315
A:Molecule type: mRNA
A:Residues: 191-397 <BER2>
C:Genetics:
A:Introns: 248/1; 300/2
C:Superfamily: plant chitinase homology
C:Keywords: glycosidase; hydrolase
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-439/Product: chitinase #status predicted <MAT>
F:183-423/Domain: plant chitinase homology <PCH>

Query Match          57.8%; Score 52; DB 2; Length 439;
Best Local Similarity 69.2%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RRRPPYLRPRPP 15
DB 50 RPTPRPRTPRPRPP 62

RESULT 18
S71169
Protein kinase, 54K (EC 2.7.1.-) - Arabidopsis thaliana
N:Alternate names: protein F4P12.270
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S71169; T45897
R:Kunomori, T.; Yamamoto, M.
submitted to the EMBL Data Library, January 1995
A:Description: A.thaliana genes encoding protein kinases of a new family.
A:Reference number: S71169
A:Accession: S71169
A:Molecule type: mRNA
A:Residues: 1-467 <KUR>
A:Cross-references: EMBL:D45354; NID:g642131; PID:g642132
R:Blöcker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Queller, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23016
A:Accession: T45897
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <BLO>
A:Cross-references: EMBL:AL132966
A:Experimental source: cultivar Columbia; BAC clone F4P12
C:Genetics:
A:Map position: 3
A:Introns: 56/3; 115/1; 124/1; 175/2; 191/3; 230/1; 260/3; 313/1; 334/3; 367/2; 402/3
A:Note: F4P12.270
C:Superfamily: fission yeast probable kinase; protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase
F:113-401/Domain: protein kinase homology <KIN>

Query Match          57.8%; Score 52; DB 2; Length 467;
Best Local Similarity 52.4%; Pred. No. 22;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

QY 1 RRRPR-----PYLRPRPP 15
DB 35 RKRPRUTWDAAPLPPPPP 55

RESULT 19
G86292
hypothetical protein AAF82153.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86292
```

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hultzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Malt, R.; Marshall, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, K.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719
 A:Accession: G86292
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1006 <STO>
 A:Cross-references: GB:AE005172; NID:98927662; PIDN:AAE82153.1; GSPDB:GN00141
 A:Genetics:
 A:Map position: 1

Query Match
 Best Local Similarity 57.8%; Score 52; DB 2; Length 1006;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 RPPPYLPFRPP 14
 DB 99 RPPPYLPFRPP 109

RESULT 20
 JC4155
 protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 14 - human
 N:Alternate names: PEZ protein; protein-tyrosine-phosphatase/ezrin-like protein
 C:Species: Homo sapiens (man)
 C>Date: 27-Aug-1995 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000
 C:Accession: JC4155
 R:Smith, A.L.; Mitchell, P.J.; Shipley, J.; Gusterson, B.A.; Rogers, M.V.; Crompton, M.F.
 Biochem. Biophys. Res. Commun. 209, 959-965, 1995
 A:Title: PEZ: a novel human cDNA encoding protein tyrosine phosphatase and ezrin-like dom
 A:Reference number: JC4155; MUID:95251727
 A:Accession: JC4155
 A:Molecule type: mRNA
 A:Residues: 1-1187 <SMI>
 A:Cross-references: EMBL:X82676; NID:93292753; PIDN:CAAF57993.1; PID:9809029
 A:Experimental source: breast
 C:Genetics:
 A:Gene: GDB:PPP14
 A:Cross-references: GDB:454485
 A:Map position: 1q32.2-1q32.2
 C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 14; protein 4.1 membrane-t
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
 F:23-302/Domain: protein 4.1 membrane-binding domain homology <BA1>
 F:566-575/Region: proline-rich
 F:709-716/Region: acidic
 F:933-1169/Domain: protein-tyrosine-phosphatase homology <PRP2>
 F:1121/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1127/Binding site: substrate phosphate (Arg) #status predicted

Query Match
 Best Local Similarity 57.2%; Score 51.5; DB 1; Length 1187;
 Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 3 RPPPYLPFRPP 14
 DB 565 RPPPYLPFRPP 575

RESULT 21
 JC2366
 protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 14 - mouse
 N:Alternate names: PEZ protein; protein-tyrosine-phosphatase PRP36; protein-tyrosine-pho
 C:Species: Mus musculus (house mouse)

C>Date: 24-Feb-1995 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000
 C:Accession: JC2366
 R:Sawada, M.; Ogata, M.; Fujino, Y.; Hamaoka, T.
 Biochem. Biophys. Res. Commun. 203, 479-484, 1994
 A:Title: cDNA cloning of a novel protein tyrosine phosphatase with homology to cytosk
 A:Reference number: JC2366; MUID:94354845
 A:Accession: JC2366
 A:Molecule type: mRNA
 A:Residues: 1-1189 <SAW>
 A:Cross-references: GB:DJ1842; NID:9507330; PIDN:BAA06628.1; PID:9507331
 A:Experimental source: thymus
 C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 14; protein 4.1 membran
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphat
 F:23-302/Domain: protein 4.1 membrane-binding domain homology <BA1>
 F:566-575/Region: proline-rich
 F:712-718/Region: acidic
 F:935-1171/Domain: protein-tyrosine-phosphatase homology <PRP2>
 F:1123/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1129/Binding site: substrate phosphate (Arg) #status predicted

Query Match
 Best Local Similarity 57.2%; Score 51.5; DB 1; Length 1189;
 Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 3 RPPPYLPFRPP 14
 DB 565 RPPPYLPFRPP 575

RESULT 22
 S43791
 PBDX protein - human
 C:Species: Homo sapiens (man)
 C>Date: 13-Jan-1995 #sequence_revision 10-Nov-1995 #text_change 07-May-1999
 C:Accession: S43791
 R:Ellis, N.A.; Ye, T.Z.; Patton, S.; German, J.; Goodfellow, P.N.; Weller, P.
 Nature Genet. 6, 394-399, 1994
 A:Title: Cloning of PBDX, an MIC2-related gene that spans the pseudautosomal boundar
 A:Reference number: S43791; MUID:94332149
 A:Accession: S43791
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-180 <ELL>
 A:Cross-references: GB:S73261; NID:9639633; PID:9639634

Query Match
 Best Local Similarity 56.7%; Score 51; DB 2; Length 180;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 3 RPPPYLPFRPP 15
 DB 47 RPPPYLPFRPP 59

RESULT 23
 JW0105
 synaptotagmin 2 alpha protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 08-Oct-1999
 C:Accession: JW0105
 R:Seet, L.F.; Cho, S.; Hessel, A.; Dumont, D.J.
 Biochem. Biophys. Res. Commun. 247, 116-122, 1998
 A:Title: Molecular cloning of multiple isoforms of synaptotagmin 2 and assignment of t
 A:Reference number: JW0105; MUID:98300294
 A:Contents: Liver
 A:Accession: JW0105
 A:Molecule type: mRNA
 A:Residues: 1-1216 <SEE>
 A:Cross-references: GB:AF041862; NID:93241994; PIDN:AAC40146.1; PID:93241995

Query Match 56.1%; Score 50.5; DB 2; Length 1216;
 Best Local Similarity 71.4%; Pred. No.: 86;
 Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 3 RRRPPYLP-RRPP 15
 DB 1014 RRRPPHPGRRPP 1027

RESULT 24

S/6867
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S76867
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 A:Reference number: S74322; MUID:97061201
 A:Accession: S76867
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <KAN>
 A:Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BA18779.1; PID:g165386
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 55.6%; Score 50; DB 2; Length 192;
 Best Local Similarity 69.2%; Pred. No.: 16;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RRRPPYLP-RRPP 14
 DB 3 RRRPPYLP-RRPP 15

RESULT 25

T06753
 zinc finger protein F15B8.140 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
 C:Accession: T06753
 R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Salanoubat, M.; Mewes,
 submitted to the Protein Sequence Database, April 1999

A:Reference number: T15794
 A:Accession: T06753
 A:Molecule type: DNA
 A:Residues: 1-383 <QUE>
 A:Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.140
 A:Experimental source: cultivar Columbia; BAC clone F15B8
 C:Genetics:
 A:Gene: ATSP:F15B8.140
 A:Map position: 3
 A:Introns: 233/3

Query Match 55.6%; Score 50; DB 2; Length 383;
 Best Local Similarity 64.3%; Pred. No.: 32;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2 RRRPPYLP-RRPP 15
 DB 52 RRRPPYLP-RRPP 65

RESULT 26

A34170
 acrosin (EC 3.4.21.10) precursor - pig
 N:Alternate names: 53K fucose-binding protein
 C:Species: *Sus scrofa domestica* (domestic pig)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34170; S08994; S02428; S04940; S16657; S02780; S10695; S12968
 R:Baba, T.; Kashiwabara, S.; Watanabe, K.; Itoh, H.; Michikawa, Y.; Kimura, K.; Takad
 J. Biol. Chem. 264, 11920-11927, 1999
 A:Title: Activation and maturation mechanisms of boar acrosin zymogen based on the de
 A:Reference number: A34170; MUID:89308595
 A:Accession: A34170

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-415 <BAB>
 A:Cross-references: GB:J04950; NID:g164702; PIDN:AAA31131.1; PID:g164703
 R:Ceohova, D.; Toepfer-Petersen, E.; Zucker, A.; Jonakova, V.
 Biol. Chem. Hoppe-Seyler 371, 317-323, 1990
 A:Title: Is spermogen a modified proacrosin? Isolation, purification, and partial c
 A:Reference number: S08994; MUID:90253655
 A:Accession: S08994

A:Molecule type: protein
 A:Residues: 'X', 18, 'X', 20-25, 'X', 27-32, 'X', 34-38, 'X', 40-50 <CEC>
 R:Toepfer-Petersen, E.; Henschen, A.
 FEBS Lett. 226, 38-42, 1987
 A:Title: Acrosin shows zona and fucose binding, novel properties for a serine protein
 A:Reference number: S02428; MUID:88083633
 A:Accession: S02428

A:Molecule type: protein
 A:Residues: 17-32; 40-55 <TOE>
 R:Adham, I.M.; Klemm, U.; Maier, W.M.; Hoyer-Pender, S.; Tsousidou, S.; Engel, W.
 Eur. J. Biochem. 182, 563-568, 1989
 A:Title: Molecular cloning of preproacrosin and analysis of its expression pattern in
 A:Reference number: S04940; MUID:89325301
 A:Accession: S04940

A:Molecule type: mRNA
 A:Residues: 1-7, 9-210, 'Q', 212-216, 'VT', 219-346, 'A', 348-388, 390-398, 'KELL', 399
 A:Cross-references: EMBL:X14844
 A:Note: the difference at the carboxyl end is due to a frameshift error
 R:Baba, T.; Michikawa, Y.; Kawakura, K.; Arai, Y.
 FEBS Lett. 244, 132-136, 1989
 A:Title: Activation of boar proacrosin is effected by processing at both N- and C-ter
 A:Reference number: S02780; MUID:89171246
 A:Accession: S02780

A:Molecule type: protein
 A:Residues: 17-69 <BA2>
 R:Toepfer-Petersen, E.; Steinberger, M.; von Eschenbach, C.E.; Zucker, A.
 FEBS Lett. 265, 51-54, 1990
 A:Title: Zona pellucida-binding of boar sperm acrosin is associated with the N-termin
 A:Reference number: S10695; MUID:90306316
 A:Accession: S10695

A:Molecule type: protein
 A:Residues: 40-62 <TO2>
 R:Toepfer-Petersen, E.; Calvete, J.; Schaefer, W.; Henschen, A.
 FEBS Lett. 275, 139-142, 1990
 A:Title: Complete localization of the disulfide bridges and glycosylation sites in bo
 A:Reference number: S12968; MUID:91085546
 A:Accession: S12968

A:Molecule type: protein
 A:Residues: 17-29; 34-66; 68-91; 94-123; 123-166; 171-184; 190-207; 209-216; 219-228; 231-245;
 C:Superfamily: acrosin; trypsin homology
 C:Keywords: glycoprotein; hydrolase; serine proteinase; sperm
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-415/Product: acrosin #status experimental <MAT>
 F:17-39/Product: acrosin light (A) chain #status experimental <LCH>
 F:40-415/Product: acrosin heavy (B) chain #status experimental <HCH>
 F:40-283/Domain: trypsin homology <TRY>
 F:300-374/Region: proline-rich
 F:19,208/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:22-152,26-160,71-87,175-244,207-223,234-264/Disulfide bonds: #status experimental

F:86,140,238/Active site: His, Asp, Ser #status predicted

Query Match

Best Local Similarity 55.6%; Score 50; DB 1; Length 415;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 RRPPLPRPP 15
| | | | | | | | | |
Db 338 RRPPLPRPP 350

RESULT 27

S29599

acrosin (EC 3.4.21.10) precursor - guinea pig (fragment)

C:Species: Cavia porcellus (guinea pig)

C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999

C:Accession: S29599

R:Gerton, G.L.; Hoff, H.B.; Baba, T.

submitted to the EMBL Data Library, May 1992

A:Description: The amino acid sequence of guinea pig proacrosin deduced from its cDNA se

A:Reference number: S29599

A:Accession: S29599

A:Molecule type: mRNA

A:Residues: 1-421 <GER>

A:Cross-references: EMBL:Z12153; NID:g49559; PIDN:CAA78137.1; PID:g49560

C:Superfamily: acrosin; trypsin homology

C:Keywords: glycoprotein; hydrolase; serine proteinase

F:41-284/Domain: trypsin homology <TRY>

Query Match

Best Local Similarity 55.6%; Score 50; DB 2; Length 421;
Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 2 RRPPLPRPP 15
| | | | | | | | | |
Db 339 RRPPLPRPP 354

RESULT 28

A54964

spliceosome-associated protein SAP-49 - human

C:Species: Homo sapiens (man)

C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 24-Sep-1999

C:Accession: A54964

R:Champion-Arnaud, P.; Reed, R.

Genes Dev. 8, 1974-1983, 1994

A:Title: The prespliceosome components SAP 49 and SAP 145 interact in a complex implicat

A:Reference number: A54964; MUID:95047348

A:Accession: A54964

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-424 <CHA>

A:Cross-references: GB:L35013; NID:9556216; PIDN:AAA60300.1; PID:9556217

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotei

F:14-81/Domain: ribonucleoprotein repeat homology <RMA4>

F:101-169/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 55.6%; Score 50; DB 2; Length 424;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRPPLPRPP 15
| | | | | | | | | |
Db 397 RRPPLPRPP 411

RESULT 29

S16748

proline-rich protein - rape (fragment)

C:Species: Brassica napus (rape)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Sep-1997

C:Accession: S16748

R:Roberts, M.R.; Foster, G.D.; Blundell, R.P.; Robinson, S.W.; Draper, J.; Scott, R.J

submitted to the EMBL Data Library, August 1991

A:Description: Cloning and characterization of a proline-rich gene expressed specific

A:Reference number: S16748

A:Accession: S16748

A:Molecule type: mRNA

A:Residues: 1-449 <ROB>

A:Cross-references: EMBL:X60376; NID:g22596; PID:g22597

Query Match

Best Local Similarity 55.6%; Score 50; DB 2; Length 449;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 RRPPLPRPP 15
| | | | | | | | | |
Db 5 QRPPLPRPP 17

RESULT 30

C96828

unknown protein F19K16.29 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96828

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

Chin, C.W.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitly, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: C96828

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-547 <STO>

A:Cross-references: GB:AE005173; NID:g6453869; PIDN:AAF09053.1; GSPDB:GNO0141

C:Genetics:

A:Gene: F19K16.29

A:Map position: 1

Query Match

Best Local Similarity 55.6%; Score 50; DB 2; Length 547;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 RRPPLPRPP 15
| | | | | | | | | |
Db 5 RRPPLPRPP 15

RESULT 31

E08B1F

immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)

C:Species: suid herpesvirus 1

C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 16-Feb-1997

C:Accession: S04713

R:Cheung, A.K.

Nucleic Acids Res. 17, 4637-4646, 1989

A:Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies

A:Reference number: S04713; MUID:89315207

A:Accession: S04713

A:Molecule type: DNA

A:Residues: 1-1460 <CHE>

C:Superfamily: herpesvirus immediate-early protein IE175

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 55.6%; Score 50; DB 1; Length 1460;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRRPPPYLPPRP 15
|:|:| | | | |
DB 156 RRRPPPYLPPRP 170

RESULT 32
T18995
hypothetical protein c06b8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18995
R:Steward, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19057
A:Accession: T18995
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3036 <MIL>
A:Cross-references: EMBL:Z81463; PIDN:CAB03852.1; GSPDB:GN00023; CESP:C06B8.7
C:Genetics:
A:Gene: CESP:C06B8.7
A:Map position: 5
A:Introns: 102/3; 157/2; 288/2; 439/3; 542/1; 591/3; 798/3; 835/3; 1139/1; 1204/3; 1325/

Query Match 55.6%; Score 50; DB 2; Length 3036;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRRPPPYLPPRP 15
|:|:| | | | |
DB 2965 RRRPPPYLPPRP 2979

RESULT 33
T29299
hypothetical protein C50F7.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T29299
R:Johnson, D.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C50F7.
A:Reference number: Z20601
A:Accession: T29299
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-589 <JOH>
A:Cross-references: EMBL:U41557; PIDN:AAA83307.1; CESP:C50F7.2
C:Genetics:
A:Gene: CESP:C50F7.2
A:Introns: 12/2
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology

Query Match 55.0%; Score 49.5; DB 2; Length 589;
Best Local Similarity 73.3%; Pred. No. 56;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2 RRRPPPYLPPRP 15
|:|:| | | | |
DB 105 RRRPPPYLPPRP 119

RESULT 34
T19345
hypothetical protein C17D12.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19345
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19112
A:Accession: T19345
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-118 <WIL>
A:Cross-references: EMBL:Z81473; PIDN:CAB03893.1; GSPDB:GN00019; CESP:C17D12.5
C:Genetics:
A:Gene: CESP:C17D12.5
A:Map position: 1
A:Introns: 22/3; 74/1

Query Match 54.4%; Score 49; DB 2; Length 118;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 PRPPYLPPRP 15
|:|:| | | | |
DB 89 PRPPYLPPRP 100

RESULT 35
JC5572
proline-rich protein P-BI precursor - human
C:Species: Homo sapiens (man)
C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
C:Accession: JC5572
R:Isemura, S.; Saitoh, E.
J. Biochem. 121, 1025-1030, 1997
A:Title: Nucleotide sequence of gene PBI encoding a protein homologous to salivary pr
A:Reference number: JC5572; MUID:98014462
A:Accession: JC5572
A:Molecule type: DNA
A:Residues: 1-134 <ISE>
A:Cross-references: DBJ:D89501; NID:q1854451; PIDN:BA13971.1; PID:q1854452
C:Comment: This protein plays roles in the maintenance of teeth by inhibiting crystal
tion from the digestive organs.
C:Genetics:
A:Introns: 18/3
C:Superfamily: proline-rich peptide P-B
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-134/Product: proline-rich protein P-BI #status predicted <MAT>
F:23-45/Domain: amino-terminal propeptide #status predicted <ATP>
F:46-58,59-71,72-84,85-97,98-112/Region: repeat
F:113-134/Domain: carboxyl-terminal peptide #status predicted <CTP>

Query Match 54.4%; Score 49; DB 2; Length 134;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PRPPYLPPRP 15
|:|:| | | | |
DB 103 PRPPYLPPRP 114

RESULT 36
F72593
hypothetical protein APE1215 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72593
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339

A:Accession: F72593
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <KAM>
 A:Cross-references: DBJ:AP000061; NID:g5104821; PIDN:BAA80204.1; PID:dl043990; PID:g510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1215

Query Match 54.4%; Score 49; DB 2; Length 161;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRPPPYLPPRRP 15
 | | | | |
 Db 144 RLGSPPLQPPRRPP 158

RESULT 37

T33700
 hypothetical protein F49F1.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

R:Miller, N.; Wamsley, P.
 Submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F49F1.

A:Reference number: 221389

A:Accession: T33700

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-210 <M1D>

A:Cross-references: EMBL:AF100656; PIDN:AAC68952.1; GSPDB:GN00022; CESP:F49F1.9

A:Experimental source: strain Bristol N2; clone F49F1

C:Genetics:

A:Gene: CESP:F49F1.9

A:Map position: 4

A:Introns: 130/3

Query Match 54.4%; Score 49; DB 2; Length 210;
 Best Local Similarity 69.2%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRRPPPYLPPRRP 14
 | | | | |
 Db 53 RPPRPPPPRRPP 65

RESULT 38

T22261
 hypothetical protein F46A8.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

R:Accession: T22261

C:Genetics:

A:Reference number: 219539

A:Accession: T22261

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-218 <M1D>

A:Cross-references: EMBL:Z81539; PIDN:CAB04391.1; GSPDB:GN00019; CESP:F46A8.5

A:Experimental source: clone F46A8

C:Genetics:

A:Gene: CESP:F46A8.5

A:Map position: 1

A:Introns: 138/3

Query Match 54.4%; Score 49; DB 2; Length 218;
 Best Local Similarity 69.2%; Pred. No. 24;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRRPPPYLPPRRP 14
 | | | | |
 Db 61 RPPRPPPPRRPP 73

RESULT 39

A27319
 gliadin - wheat

C:Species: Triticum aestivum (common wheat)
 C:Date: 04-Mar-1988 #sequence_revision 04-Mar-1988 #text_change 03-Feb-1994

R:Accession: A27319

R:Reeves, C.D.; Okita, T.W.

Gene 52, 257-266, 1987

A:Title: Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats.

A:Reference number: A27319; MUID:87277398

A:Accession: A27319

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-296 <REE>

C:Superfamily: gliadin

Query Match 54.4%; Score 49; DB 2; Length 296;
 Best Local Similarity 61.5%; Pred. No. 32;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPPPQLPPYQPP 15
 | | | | |
 Db 85 QPFPQQLPPYQPP 97

RESULT 40

S07361
 alpha/beta-gliadin precursor (clone PM1215) - wheat

C:Species: Triticum aestivum (common wheat)
 C:Date: 08-Jun-1994 #sequence_revision 01-Dec-1995 #text_change 20-Aug-1999

C:Accession: S07361

R:Summer-Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Soell, D.

Nucleic Acids Res. 13, 3905-3916, 1985

A:Title: Conservation and variability of wheat alpha/beta-gliadin genes.

A:Reference number: S07361; MUID:85242077

A:Accession: S07361

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-296 <SUM>

A:Cross-references: EMBL:X02538; NID:g21756; PIDN:CAA26383.1; PID:g21757

C:Superfamily: gliadin

C:Keywords: seed; storage protein

Query Match 54.4%; Score 49; DB 2; Length 296;
 Best Local Similarity 61.5%; Pred. No. 32;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPPPQLPPYQPP 15
 | | | | |
 Db 85 QPFPQQLPPYQPP 97

RESULT 41

F84799
 hypothetical protein At2g37990 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: F84799

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: F84799
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <STO>
A:Cross-references: GB:AE002093; NID:g4895185; PIDN:AA032772.1; GSPDB:GN00139
C:Genetics:
A:Gene: ATg37990
A:Map position: 2

Query Match 54.4%; Score 49; DB 2; Length 352;
Best Local Similarity 58.3%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 PRPPYLRPRRP 15
| | | | | | | | | |
Db 80 PRKHPRPRRP 91

RESULT 42
320500
hydroxyproline-rich glycoprotein - rice
C:Species: Oryza sativa (rice)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S20500
R:Caellies, C.; Delsen, M.; Puigdomenech, P.
Plant Mol. Biol. 18, 617-619, 1992
A:title: The hydroxyproline-rich glycoprotein gene from Oryza sativa.
A:Reference number: S20500; MUID:92163031
A:Accession: S20500
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-369 <CAE>
A:Cross-references: EMBL:X61280; NID:g433815; PIDN:CAA43583.1; PID:g433816
A:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein

Query Match 54.4%; Score 49; DB 2; Length 369;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 PPTLRPRRP 15
| | | | | | | | | |
Db 94 PPTPKRPRRP 103

RESULT 43
T32944
hypothetical protein W01B11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32944
R:Bradshaw, H.; Graves, T.; Blair, T.
submitted to the EMBL Data Library, January 1998
A:Description: The sequence of C. elegans cosmid W01B11.
A:Reference number: Z21250
A:Accession: T32944
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <BRA>
A:Cross-references: EMBL:AF043704; PIDN:AA97600.1; GSPDB:GN00019; CESP:W01B11.5
A:Experimental source: strain Bristol N2; clone W01B11
C:Genetics:
A:Gene: CESP:W01B11.5
A:Map position: 1
A:introns: 27/3; 44/3; 62/2; 241/2; 270/1; 293/1

Query Match 54.4%; Score 49; DB 2; Length 380;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 RRRPPYLRPRRP 15
| | | | | | | | | |
Db 121 RRRPPALSPTRRP 135

RESULT 44
E83802
hypothetical protein BH1221 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: E83802
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20263314
A:Accession: E83802
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04940.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1221

Query Match 54.4%; Score 49; DB 2; Length 444;
Best Local Similarity 53.8%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRRPPYLRPRRP 15
| | | | | | | | | |
Db 121 QRPVAPTRPRRP 133

RESULT 45
T21275
hypothetical protein F22E12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21275
R:Mathews, P.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19399
A:Accession: T21275
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1203 <WTL>
A:Cross-references: EMBL:Z71180; PIDN:CAA94892.1; GSPDB:GN00023; CESP:F22E12.1
C:Genetics:
A:Gene: CESP:F22E12.1
A:Map position: 5
A:introns: 71/1; 156/3; 223/3; 339/1; 669/3; 736/1; 758/3; 854/3; 931/3; 968/3; 1181/

Query Match 54.4%; Score 49; DB 2; Length 1203;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 RRRPPYLRPRRP 15
| | | | | | | | | |
Db 707 RRRPPQTRPRRP 720

Search completed: September 24, 2001, 10:06:03
job time: 43 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:05:20 ; Search time 16.18 Seconds
(without alignments)
31.757 Million cell updates/sec

Title: US-09-276-868-3
Perfect score: 90
Sequence: 1 RRRPRPYLPRRPP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	172	PR39_PIG	P80054 sus scrofa
2	66	73.3	190	BC77_BOVIN	P19661 bos taurus
3	59.5	66.1	107	COT7_BACSU	P18653 bacillus su
4	58	64.4	190	BC77_SHEEP	P50415 ovis aries
5	56.5	62.8	168	AP14_APIME	O06601 apis mellif
6	55.5	61.7	151	RNB_HSV2H	P89479 herpes simp
7	53	58.9	144	AP22_APIME	P35581 apis mellif
8	52	57.8	283	AP73_APIME	O06602 apis mellif
9	52	57.8	261	RL1_HSV2H	P28283 herpes simp
10	52	57.8	354	ATH1_HUMAN	O92858 homo sapien
11	52	57.8	467	AECL1_ARATH	P51566 arabidopsis
12	52	57.8	841	RELA_STRAT	O85709 streptomyc
13	51.5	57.2	1187	PTNE_HUMAN	O15678 homo sapien
14	51.5	57.2	1189	PTNE_MOUSE	O62130 mus musculu
15	51	56.7	180	MX1_PALPR	P80408 palomera pr
16	51	56.7	151	XG_HUMAN	P55808 homo sapien
17	50.5	56.1	393	CIW4_HUMAN	O99498 homo sapien
18	50	55.6	17	AP1D_BOMRA	P81464 bombus pasc
19	50	55.6	415	ACRO_PIG	P08001 sus scrofa
20	50	55.6	424	SP49_HUMAN	O15427 homo sapien
21	50	55.6	449	APG_BRANA	P40603 brassica na
22	50	55.6	678	ABBP_RIPCL	O27905 tripturus c
23	49	54.4	296	GDA6_WHEAT	P04726 triticum ae
24	49	54.4	2911	FBH2_HUMAN	P35556 homo sapien
25	48.5	53.9	2142	BAV12_HUMAN	P48534 homo sapien
26	48	53.3	281	FASL_HUMAN	P48023 homo sapien
27	48	53.3	402	VGLD_PVRV1	P07445 pseudorabie
28	48	53.3	520	CP84_ARATH	O42600 arabidopsis
29	48	53.3	759	TOP3_CAEEL	O61660 caenorhabdi
30	47.5	52.8	283	EXTN1_SORBI	P24152 sorghum dlc
31	47.5	52.8	372	DBPA_HUMAN	P16589 homo sapien
32	47.5	52.8	1443	SVJ2_HUMAN	O15056 homo sapien
33	47	52.2	176	BC75_BOVIN	P19660 bos taurus

34	47	52.2	261	1	PRP2_MOUSE	P05142 mus musculu
35	47	52.2	296	1	PRP3_MOUSE	P05143 mus musculu
36	47	52.2	507	1	MEFA_HUMAN	O02078 homo sapien
37	47	52.2	534	1	APG_ARATH	P40602 arabidopsis
38	47	52.2	846	1	IRSL_HCMVA	P09715 human cytom
39	47	52.2	1040	1	BO12_YEAST	P39969 saccharomyc
40	47	52.2	1085	1	CASR_BOVIN	P35384 bos taurus
41	46.5	51.7	267	1	EXTN1_MOUSE	P14918 zea mays (m
42	46.5	51.7	424	1	SP60_DICDI	P15270 dictyostell
43	46.5	51.7	1245	1	POL5_SINDV	P03316 sindbis vir
44	46	51.1	138	1	Y028_NPOVP	O10295 oryza pseu
45	46	51.1	176	1	C561_ECOLI	P08732 escherichia

ALIGNMENTS

RESULT 1
PR39_PIG STANDARD; PRT; 172 AA.
AC P80054; Q9TR84;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ANTI-BACTERIAL PROTEIN PR-39 PRECURSOR.
GN PR39.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RX MEDLINE=95350216; PubMed=7624374;
RA Gudmundsson G.H., Magnusson K.P., Chowdhary B.P., Johansson M.,
RA Andersson L., Boman H.G.;
RT "Structure of the gene for porcine peptide antibiotic PR-39, a
RT cathelin gene family member: comparative mapping of the locus for the
RT human peptide antibiotic FALL-39.";
RT Proc. Natl. Acad. Sci. U.S.A. 92:7085-7089(1995).
RN [2]
RX MEDLINE=96105365; PubMed=7498526;
RA Zhou C., Ganz T., Lehrer R.I.;
RT "Structures of genes for two cathelin-associated antimicrobial
RT peptides: prophenin-2 and PR-39.";
RT FEBS Lett. 376:130-134(1995).
RN [3]
RX MEDLINE=96105365; PubMed=7498526;
RA Zhou C., Ganz T., Lehrer R.I.;
RT "Structures of genes for two cathelin-associated antimicrobial
RT peptides: prophenin-2 and PR-39.";
RT FEBS Lett. 376:130-134(1995).
RN [4]
RX MEDLINE=92111534; PubMed=1765098;
RA Ageberth B., Lee J.-Y., Bergman T., Carlquist M., Boman H.G.,
RT "Amino acid sequence of PR-39. Isolation from pig intestine of a new
RT member of the family of proline-arginine-rich antibacterial
RT peptides.";
RT Eur. J. Biochem. 202:849-854(1991).
RN [5]
RX MEDLINE=95088504; PubMed=7996056;
RA Shi J., Ross C.R., Chengappa M.M., Blecha F.;
RT "Identification of a proline-arginine-rich antibacterial peptide from
RT neutrophils that is analogous to PR-39, an antibacterial peptide from
RT the small intestine.";

RL J. Leukoc. Biol. 56:807-811(1994).
 CC -1- FUNCTION: EXERTS A POTENT ANTIMICROBIAL ACTIVITY AGAINST BOTH
 CC E. COLI AND B. MEGATERIUM.
 CC -1- TISSUE SPECIFICITY: SMALL INTESTINE AND BONE MARROW.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X87236; CAA60682.1; -
 CC EMBL; L23825; AAA31109.1; -
 CC EMBL; X89201; CAA61487.1; -
 CC PIR; S19563; S19563.
 CC InterPro; IPR001894; -
 CC Pfam; PF00666; Cathelicidins; 1.
 CC DR PROSITE; PS00946; CATHELICIDINS_1; 1.
 CC DR PROSITE; PS00947; CATHELICIDINS_2; 1.
 CC KW Antibiotic; Amidation; Signal.
 CC FT SIGNAL 1 29
 CC FT PROPEP 30 130
 CC FT CHAIN 131 169
 CC FT MOD_RES 30 30
 CC FT MOD_RES 30 30
 CC FT DISULFID 85 96
 CC FT DISULFID 107 124
 CC FT MOD_RES 169 169
 CC FT CONFLICT 21 21
 CC FT CONFLICT 29 29
 CC FT CONFLICT 90 91
 CC FT CONFLICT 117 119
 CC FT CONFLICT 157 157
 CC SQ SEQUENCE 172 AA; 19476 MW; 994B92798C0E133 CRC64;
 P -> I (IN REF. 5).
 IHS -> NDP (IN REF. 1).
 ANTIMICROBIAL PROTEIN PR-39.
 PYRROLIDONE CARBOXYLIC ACID (BY
 SIMILARITY).
 BY SIMILARITY.
 AMIDATION (G-170 PROVIDE AMIDE GROUP).
 G -> A (IN REF. 2).
 A -> T (IN REF. 1).
 RQ -> OR (IN REF. 1).
 IHS -> NDP (IN REF. 1).
 P -> I (IN REF. 5).
 Query Match 100.0%; Score 90; DB 1; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA Scocchi M., Wang S., Zanetti M.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE OF 131-189.
 RC TISSUE=Neutrophils;
 RX MEDLINE=91035404; PubMed=2229048;
 RA Frank R.W., Gennaro R., Schneider K., Przybylski M., Romeo D.;
 RT "Amino acid sequences of two proline-rich bactericins. Antimicrobial
 RL peptides of bovine neutrophils.";
 RL J. Biol. Chem. 265:18871-18874(1990).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=96300243; PubMed=8706679;
 RA Storici P., Tossi A., Lenarcic B., Romeo D.;
 RT "Purification and structural characterization of bovine
 RT cathelicidins, precursors of antimicrobial peptides.";
 RL Eur. J. Biochem. 238:769-776(1996).
 CC -1- FUNCTION: EXERTS, IN VITRO, A POTENT ANTIMICROBIAL ACTIVITY.
 CC PROBABLY DUE TO AN IMPAIRMENT OF THE FUNCTION OF THE RESPIRATORY
 CC CHAIN AND OF ENERGY-DEPENDENT ACTIVITIES IN THE INNER MEMBRANE
 CC OF SUSCEPTIBLE MICROORGANISMS.
 CC -1- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
 CC -1- PM: ELASTASE IS RESPONSIBLE FOR ITS MATURATION.
 CC -1- MASS SPECTROMETRY: MW=18395; MW_ERR=1; METHOD=ELECTROSPRAY;
 CC RANGE=30-190.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L42977; AA87359.1; -
 CC EMBL; Y09471; CAA70616.1; -
 CC PIR; A36589; A36589.
 CC InterPro; IPR001894; -
 CC Pfam; PF00666; Cathelicidins; 1.
 CC DR PROSITE; PS00946; CATHELICIDINS_1; 1.
 CC DR PROSITE; PS00947; CATHELICIDINS_2; 1.
 CC KW Antibiotic; Repeat; Signal.
 CC FT SIGNAL 1 29
 CC FT PROPEP 30 130
 CC FT CHAIN 131 190
 CC FT PROPEP 189 190
 CC FT MOD_RES 30 30
 CC FT DISULFID 85 96
 CC FT DISULFID 107 124
 CC SQ SEQUENCE 190 AA; 21567 MW; 8CD07D7AA30A731C CRC64;
 BACTENECIN 7.
 REMOVED PARTIALLY.
 PYRROLIDONE CARBOXYLIC ACID.
 Query Match 73.3%; Score 66; DB 1; Length 190;
 Best Local Similarity 85.7%; Pred. No. 0.12;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OC Bacillus/Staphylococcus group: Bacillus.
OX NCBI_TaxID=1423;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=89313296; PubMed=2546006;
RA Aronson A.I., Song H.Y., Bourne N.;
RT "Gene structure and precursor processing of a novel Bacillus subtilis
  spore coat protein." 3:437-444(1989).
RL MOL. Microbiol. 3:437-444(1989).
CC -1- FUNCTION: POSSIBLY PROTECTION OF SPORE AND PROBABLY PLAYS
  SOME ROLE IN GERMINATION.
CC -1- SUBCELLULAR LOCATION: OUTER SURFACE OF ENDOSPORE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: X13740; CA32004.1; -
ER EMBL: Z99110; CAB13066.1; -
ER PIR: S04835; S04835.
DR Subtilist; BG10495; cdt.
KW Sporulation; Signal.
FT SIGNAL 1 44
FT CHAIN 45 107 SPORE COAT PROTEIN T.
SQ SEQUENCE 107 AA; 12992 MW; AD1F66F0C4CE29A3 CRC64;

Query Match 66.1%; Score 59.5; DB 1; Length 107;
Best Local Similarity 84.6%; Pred. No. 0.39;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 4 PRPP-YLPRPP 15
ID BCT7-SHEEP STANDARD; PRT; 190 AA.
AC P50415;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BACTENECEIN 7 PRECURSOR (BAC7).
GN BAC7.5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scocchi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins.";
RL FEBS Lett. 376:225-228(1995).
CC -1- FUNCTION: EXERS, IN VITRO, A POTENT ANTIMICROBIAL ACTIVITY.
  PROBABLY DUE TO AN IMPAIRMENT OF THE FUNCTION OF THE RESPIRATORY
  CHAIN AND OF ENERGY-DEPENDENT ACTIVITIES IN THE INNER MEMBRANE
  OF SUSCEPTIBLE MICROORGANISMS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: L46852; AAA85468.1; -
DR InterPro; IPR001894; -
DR Pfam; PF00666; Cathelicidins_1;
DR PROSITE; PS00946; CATHELICIDINS_1;
DR PROSITE; PS00947; CATHELICIDINS_2;
KW Antibiotic; Repeat; Signal.
FT SIGNAL 1 29
FT PROPEP 30 130 POTENTIAL,
FT CHAIN 131 190 BY SIMILARITY.
FT MOD_RSS 30 30 BACTENECEIN 7.
FT DISULFID 85 96 PYROLIDONE CARBOXYLIC ACID
FT DISULFID 107 124 (BY SIMILARITY).
SQ SEQUENCE 190 AA; 21829 MW; E4AAFB1600E98371 CRC64;

Query Match 64.4%; Score 58; DB 1; Length 190;
Best Local Similarity 78.6%; Pred. No. 1;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPPRYLPRPP 14
ID B132 RLRPPRRLPRPP 145
DB 132 RLRPPRRLPRPP 145

RESULT 5
AP14-APIME STANDARD; PRT; 168 AA.
ID AP14-APIME
AC 006601; P11525; P11526; P11527;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE APIDACIN PRECURSOR, TYPE 14.
GN APID14.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223697; PubMed=8467807;
RA Casteels-Josson K., Capaci T., Casteels P., Tempst P.;
RT "Apidecin multipetide precursor structure: a putative mechanism for
  amplification of the insect antibacterial response.";
RL EMBO J. 12:1569-1578(1993).
GN [2]
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemolymph;
RX MEDLINE=9005446; PubMed=2676519;
RA Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;
RT "Apidecins: antibacterial peptides from honeybees.";
RL EMBO J. 8:2387-2391(1989).
CC -1- FUNCTION: APIDACINS HAVE BACTERICIDAL ACTIVITY. PREDOMINANTLY
  AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
  PROPAagation.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL: X72575; CA51167.1; -
DR PIR: S05383; S05383.
DR PIR: S06675; S06675.
DR PIR: S06676; S06676.
DR PIR: S35330; S35330.

```

DR InterPro: IPR001979; 5.
 DR Pfam: PF00807; Apidaecin; 5.
 KW Insect immunity: Antidiotin; Hemolymph; Signal; Multigene family;
 KW Cleavage on pair of basic residues; Repeat.
 FT SIGNAL 1 19
 FT PROPEP 35 42
 FT PEPTIDE 43 60
 FT PROPEP 63 70
 FT PEPTIDE 71 88
 FT PROPEP 91 98
 FT PEPTIDE 99 116
 FT PROPEP 119 124
 FT PEPTIDE 125 142
 FT PROPEP 145 150
 FT PEPTIDE 151 168
 SQ SEQUENCE 168 AA; 19380 MW; 594B931254C0A37 CRC64;

Query Match
 Best Local Similarity 62.8%; Score 56.5; DB 1; Length 168;
 Pred. No. 1.3;
 Matches 11; Conservative 2; Mismatches 2; Indels 7; Gaps 1;
 DB 117 RRPPEAPGNRRPVYIPQPRP 138

RESULT 6
 RNB_HSV2H STANDARD; PRT; 151 AA.
 AC P89479;
 DT 01-OCT-2000 (Rel. 40; Created)
 DT 01-OCT-2000 (Rel. 40; Last sequence update)
 DT 01-OCT-2000 (Rel. 40; Last annotation update)
 DE POTENTIAL RNA-BINDING PROTEIN.
 OS Herpes simplex virus (type 2 / strain HG52).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 CC NCBI_TaxID=10315;
 RN
 RA Dojan A.;
 RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: BINDS DNA AND RNA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS (BY SIMILARITY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z86099; CAB06719.1; -
 DR DNA-binding; RNA-binding; Repeat; Nuclear protein.
 FT DOMAIN 90 146
 FT REPEAT 90 146
 FT REPEAT 96 101
 FT REPEAT 102 104
 FT REPEAT 105 110
 FT REPEAT 111 116
 FT REPEAT 117 122
 FT REPEAT 123 128
 FT REPEAT 129 130
 FT REPEAT 130 134
 FT REPEAT 135 140
 FT REPEAT 141 146
 SQ SEQUENCE 151 AA; 16297 MW; FAB751F23C3DB6AE CRC64;

Query Match 61.7%; Score 55.5; DB 1; Length 151;

Best Local Similarity 73.3%; Pred. No. 1.6;
 Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 DB 127 RRPPEAPGNRRPVYIPQPRP 141

RESULT 7
 AP22_APIME STANDARD; PRT; 144 AA.
 AC P35581; P11525; P11526;
 DT 01-OCT-1989 (Rel. 12; Created)
 DT 01-JUN-1994 (Rel. 29; Last sequence update)
 DT 01-JUN-1994 (Rel. 29; Last annotation update)
 DE APIDAECIN PRECURSOR, TYPE 22.
 OS Apis mellifera (Honeybee).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 CC Apoidea; Apidae; Apis.
 CC NCBI_TaxID=7460;
 RN
 RA
 RL
 RL EMBL J. 12:1569-1578(1993).
 RP SEQUENCE (APIDAECIN IA/IB).
 RC TISSUE=Hemolymph;
 RX MEDLINE=9323697; PubMed=8467807;
 RA Casteels J., Capaci T., Casteels P., Tempst P.;
 RT "Apidaecin multiprotein precursor structure: a putative mechanism for
 RT amplification of the insect antibacterial response.";
 RL EMBL J. 12:1569-1578(1993).
 CC
 CC -1- FUNCTION: APIDAECIN HAVE BACTERICIDAL ACTIVITY: PREDOMINANTLY
 CC AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
 CC PROPAGATION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X72576; CAAS1168.1; -
 DR PIR; S05383; S05383.
 DR PIR; S06675; S06675.
 DR PIR; S35331; S35331.
 DR InterPro: IPR001979; 5.
 DR Pfam: IPR00807; Apidaecin; 4.
 KW Insect immunity: Antidiotin; Hemolymph; Signal; Multigene family;
 KW Cleavage on pair of basic residues; Repeat.
 FT SIGNAL 1 19
 FT PROPEP 35 42
 FT PEPTIDE 43 60
 FT PROPEP 63 70
 FT PEPTIDE 71 88
 FT PROPEP 91 98
 FT PEPTIDE 99 116
 FT PROPEP 119 126
 FT PEPTIDE 127 144
 SQ SEQUENCE 144 AA; 16539 MW; 6F1AD74CE77108D CRC64;

Query Match 58.9%; Score 53; DB 1; Length 144;
 Best Local Similarity 72.7%; Pred. No. 3;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 5 RRPPEAPGNRRPVYIPQPRP 15

DB 46 RPYIPQPRPP 56

RESULT 8

AP73_APIIME	STANDARD	PRT	283 AA
ID AP73_APIIME			
AC 006602; P11525; P11526;			
DT 01-OCT-1989 (Rel. 12, Created)			
DT 01-JUN-1994 (Rel. 29, Last sequence update)			
DT 01-JUN-1994 (Rel. 29, Last annotation update)			
DE APIDAEIN PRECURSOR, TYPE 73 (FRAGMENT).			
OS APIDV3.			
OS Apis mellifera (Honeybee).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;			
OC Apoidea; Apidae; Apis.			
OC NCBI_TaxID=7460;			
RM [1]			
RP SEQUENCE FROM N.A.			
RA MEDLINE-9322697; PubMed-8467807;			
RA Casteels-Josson K., Capaci T., Casteels P., Tempst P.;			
RT "Apidecin multipeptide precursor structure: a putative mechanism for			
RT amplification of the insect antibacterial response.";			
RL EMBO J. 12:1569-1578(1993).			
RM [2]			
RP SEQUENCE OF APIDAEIN IA/IB.			
RC TISSUE=Hemolymph;			
RX MEDLINE-90005446; PubMed-2676519;			
RA Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;			
RT "Apidacins: antibacterial peptides from honeybees.";			
RL EMBO J. 8:2387-2391(1989).			
CC -1- FUNCTION: APIDAEIN HAVE BACTERICIDAL ACTIVITY; PREDOMINANTLY			
CC AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL			
CC PROPAGATION.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL: X72577; CA51169.1; -			
DR PIR: S05383; S05383.			
DR PIR: S06675; S06675.			
DR PIR: S35332; S35332.			
DR InterPro: IPR001979; -			
DR Pfam: PF00807; Apidecin; 9.			
KW Insect immunity; Antibiotic; Hemolymph; Signal; Multigene family;			
KM Cleavage on pair of basic residues; Repeat.			
FT NON_TER 1 1			
FT SIGNAL <1 18 POTENTIAL.			
FT PROPEP 34 41			
FT PEPTIDE 42 59 APIDAEIN IB.			
FT PROPEP 62 69			
FT PEPTIDE 70 87 APIDAEIN IB.			
FT PROPEP 90 97			
FT PEPTIDE 98 115 APIDAEIN.			
FT PROPEP 118 125			
FT PEPTIDE 126 143 APIDAEIN IB.			
FT PROPEP 146 153			
FT PEPTIDE 154 171 APIDAEIN.			
FT PROPEP 174 182			
FT PEPTIDE 183 199 APIDAEIN IB.			
FT PROPEP 202 209			
FT PEPTIDE 210 227 APIDAEIN IB.			
FT PROPEP 230 237			
FT PEPTIDE 238 255 APIDAEIN IB.			
FT PROPEP 258 265			
FT PEPTIDE 266 283			
FT PEPTIDE 283 AA; 32695 MM; 4EA5FEDECD5E142B CRC64;			

Query Match 58.9%; Score 53; DB 1; Length 283;
Best Local Similarity 72.7%; Pred. No. 5.6;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 RPYLPFRPP 15
||:|||||
DB 45 RPYIPQPRPP 55

RESULT 9

RL_HSV2H	STANDARD	PRT	261 AA
ID RL_HSV2H			
AC P28283;			
DT 01-DEC-1992 (Rel. 24, Created)			
DT 01-DEC-1992 (Rel. 24, Last sequence update)			
DT 01-NOV-1997 (Rel. 35, Last annotation update)			
DE NEUROVIRULENCE FACTOR (ICP34.5).			
GN RL1.			
OS Herpes simplex virus (type 2 / strain HG52).			
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC Alphaherpesvirinae; Simplexvirus.			
OC NCBI_TaxID=10315;			
RM [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-92113549; PubMed-1662697;			
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;			
RT "Comparative sequence analysis of the long repeat regions and			
RT adjoining parts of the long unique regions in the genomes of herpes			
RT simplex viruses types 1 and 2.";			
RL J. Gen. Virol. 72:3057-3075(1991).			
RM [2]			
RP SEQUENCE FROM N.A.			
RA Dolan A.;			
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL: D10471; BAA23428.1; -			
DR EMBL: Z86099; CAB06759.1; -			
DR EMBL: Z86099; CAB06706.1; -			
DR PIR: J01502; WMBEXE.			
KW Repeat.			
FT DOMAIN 3 12 2 X 5 AA TANDEM REPEATS OF R-R-R-G-P.			
FT REPEAT 3 7			
FT REPEAT 8 12			
FT DOMAIN 16 31 2 X 8 AA TANDEM REPEATS OF P-R-P-G-A-P-A-			
FT REPEAT 16 23 V.			
FT REPEAT 24 31			
FT REPEAT 24 31			
SO SEQUENCE 261 AA; 27908 MM; 4BBD13AF3D906D71 CRC64;			

Query Match 57.8%; Score 52; DB 1; Length 261;
Best Local Similarity 64.7%; Pred. No. 6.8;
Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 1 RRRPRP--PYLPFRPP 15
||||| | :||| |
DB 13 RRRPRGAPAVPRGAP 29

RESULT 10

ATHL_HUMAN	STANDARD	PRT	354 AA
ID ATHL_HUMAN			
AC Q92058;			
DT 15-DEC-1998 (Rel. 37, Created)			

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE ATOMAL PROTEIN HOMOLOG 1 (HELIX-LOOP-HELIX PROTEIN HATH-1).
GN HATH1 OR HATH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97026280; PubMed=8872459;
RA Ben-Arie N., McCall A.E., Berkman S., Elchele G., Bellen H.J.,
RA Zoghbi H.Y.;
RT "Evolutionary conservation of sequence and expression of the bHLH
RT protein Atomal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -1- FUNCTION: ACTIVATES E BOX-DEPENDENT TRANSCRIPTION IN COLLABORATION
CC WITH E47, BUT THE ACTIVITY IS COMPLETELY ANTAGONIZED BY THE
CC NEGATIVE REGULATOR OF NEUROGENESIS HES-1. MAY PLAY A ROLE IN THE
CC DIFFERENTIATION OF SUBSETS OF NEURAL CELLS BY ACTIVATING E BOX-
CC DEPENDENT TRANSCRIPTION (BY SIMILARITY).
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC bHLH PROTEIN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (bHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. "ATOMAL" SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U61148; AAB41305.1; -
DR MIM: 601461; -
DR InterPro: IPR001092; -
DR InterPro: IPR003015; -
DR Pfam: PF00010; HLH; 1.
DR PROSITE: PS00038; HELIX-LOOP-HELIX; FALSE NEG.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 29 38 FOLIT-PRO.
FT DNA_BIND 160 171 BASIC DOMAIN.
FT DOMAIN 172 212 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 224 228 POLY-PRO.
SQ SEQUENCE 354 AA; 38160 MW; AB12F1E917A00ABD CRC64;

Query Match 57.8%; Score 52; DB 1; Length 354;
Best Local Similarity 57.1%; Pred. No. 9.1;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPPVLRPRRP 15
DB 21 RQFQPHLPPQPPP 34

RESULT 11
AFCL ARATH STANDARD; PRT; 467 AA.
AC P51566;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROTEIN KINASE AFCL (EC 2.7.1.1.-).
OS AFCL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RX MEDLINE=95083650; PubMed=7991592;
RA Bender J., Fink G.R.;
RT "AFCL, a LAMMER kinase from Arabidopsis thaliana, activates STE12-
RT dependent processes in yeast.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12105-12109(1994).
CC -1- FUNCTION: ACTIVATOR OF YEAST TRANSCRIPTION FACTOR, STE12.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC LAMMER SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U16176; AAA57117.1; -
DR HSP, P24941; IAO1.
DR InterPro: IPR00719; -
DR InterPro: IPR002290; -
DR Pfam: PF00069; Kinase; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE NEG.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 115 443
FT NP_BIND 121 129 ATP (BY SIMILARITY).
FT BINDING 144 144 ATP (BY SIMILARITY).
FT ACT_SITE 240 240 BY SIMILARITY.
SQ SEQUENCE 467 AA; 54216 MW; 5AD739A82F490E12 CRC64;

Query Match 57.8%; Score 52; DB 1; Length 467;
Best Local Similarity 52.4%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 1;

OY 1 RRRP-----PPVLRPRRP 15
DB 35 RRRPRLTWDAAPLPPPPP 55

RESULT 12
RELA STRAT STANDARD; PRT; 841 AA.
AC O85709;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GTP PYROPHOSPHOKINASE (EC 2.7.6.5) (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE).
DE (PPGPP SYNTHETASE 1) ((P)PPGP SYNTHETASE).
OS RELA.
OS Streptomyces antibioticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMRU 3720;
RA Jones G.H., Hoyt S.;
RT "Rela is required for actinomycin production in Streptomyces
RT antibioticus";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IN EUBACTERIA PPGPP (GUANOSINE 3'-DIPHOSPHATE 5'-
CC DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT
CC COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO
CC CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE
CC FORMATION OF PPGPP WHICH IS THEN HYDROLYSED TO FORM PPGP (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + GTP = AMP + GUANOSINE 3'-DIPHOSPHATE
CC 5'-TRIPHOSPHATE.

```

```

CC -1- PATHWAY: FIRST STEP IN THE METABOLISM OF PGPP.
CC -1- SIMILARITY: BELONGS TO THE RETA / SPOT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AF072829; AAC26021.1; -.
CC DR InterPro; IPR002912; -.
CC DR Pfam; PF01842; ACT: 1.
CC DR Transferrase; Kinase.
CC DR SEQUENCE 841 AA; 93671 MW; 632A037BA4EF4C94 CRC64;
CC -----
CC Query Match 57.8%; Score 52; DB 1; Length 841;
CC Best Local Similarity 60.0%; Pred. No. 21;
CC Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
CC -----
CC QY 1 RRRPRPPYLPRPP 15
CC I:|:|:|:|:|:|
CC Db 50 RRPKAPRRPRPP 64
CC -----
CC RESULT 13
CC PTNE_HUMAN STANDARD; PRT; 1187 AA.
CC AC Q15678;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
CC DE PROTEIN-TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 14 (EC 3.1.3.48)
CC DE (PROTEIN-TYROSINE PHOSPHATASE PE2).
CC GN PTPN14 OR PE2.
CC CS Homo sapiens (Human).
CC CS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC CX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE-Breast carcinoma;
CC RX MEDLINE=95251727; PubMed=7733990;
CC RA Smith A.L., Mitchell P.J., Shipley J., Gusterson B.A., Rogers M.V.,
CC RA Crompton M.R.;
CC FT "pe2: a novel human cDNA encoding protein tyrosine phosphatase- and
CC FT ezrin-like domains.";
CC RL Biochem. Biophys. Res. Commun. 209:959-965(1995).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC -1- PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF HUMAN TISSUES
CC -1- INCLUDING KIDNEY, SKELETAL MUSCLE, LUNG AND PLACENTA.
CC -1- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,
CC -1- RADIOLIN, AND TNLIN.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X82676; CAA57993.1; -.
CC DR HSSP; P18052; 1YFO.
CC DR MTM; 603155; -.
CC DR InterPro; IPR000242; -.
CC DR InterPro; IPR000299; -.
CC DR InterPro; IPR000387; -.

```

```
DR Pfam: PF00373; Band_41; 1.  
DR pfam: PF00102; Y-phosphatase; 1.  
DR PRINTS: PR00700; PRTYPHPTASE.  
DR PRINTS: PR00935; BAND41.  
DR PROSITE: PS00660; BAND_41_1; 1.  
DR PROSITE: PS00661; BAND_41_2; 1.  
DR PROSITE: PS50057; BAND_41_3; 1.  
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.  
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTF; 1.  
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.  
DR Structural protein; Cytoskeleton; Hydrolyase.  
FW DOMAIN 75 239  
FT ACT_SITE 933 1187  
FT ACT_SITE 1121 1121  
FT DOMAIN 566 573  
FT POLY-PRO.  
FO POLY-GLU.  
SQ SEQUENCE 1187 AA; 135239 MW; 015760B75E357AE3 CRC64;  
  
Query Match 57.2%; Score 51.5; DB 1; Length 1187;  
Best Local Similarity 83.3%; Pred. No. 33;  
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
QY 3 RPRPEYLPRP 14  
Db 565 RPPPY-PRPR 575  
  
RESULT 14  
PTNE_MOUSE STANDARD; PRT; 1189 AA.  
AC Q62130;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 14 (EC 3.1.3.48)  
DE (PROTEIN-TYROSINE PHOSPHATASE PR36).  
GN PRPN14.  
GC Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CB-17-SCID; TISSUE=Thymus;  
RX MEDLINE=94354845; PubMed=8074693;  
RA Sawada M., Ogata M., Fujino Y., Hamaoka T.;  
RT "cDNA cloning of a novel protein tyrosine phosphatase with homology  
RL to cytoskeletal protein 4.1 and its expression in T-lineage cells.",  
CC Biochem. Biophys. Res. Commun. 203:479-484(1994).  
CC -1 FUNCTION: MAY BE INVOLVED IN THE REGULATION OF T CELL DEVELOPMENT.  
CC -1 CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -1 TISSUE SPECIFICITY: THYMUS; IN CELLS OF BOTH HEMATOPOIETIC AND  
CC NON-HEMATOPOIETIC ORIGINS.  
CC -1 SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,  
CC RADIXIN, AND TALIN.  
CC -1 SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).
```

DR InterPro; IPR000387; .
 DR Pfam; PF00373; Band_41; 1.
 DR Pfam; PF00102; Y-phosphatase; 1.
 DR PRINTS; PRO0700; PRTYPHPTASE.
 DR PRINTS; PRO0935; BAND41.
 DR PROSITE; PS00660; BAND_41_1; 1.
 DR PROSITE; PS00661; BAND_41_2; 1.
 DR PROSITE; PS50057; BAND_41_3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR Structural protein; Cytoskeleton; Hydrolase.
 KM DOMAIN 75 239 BAND 4.1-LIKE.
 FT DOMAIN 935 1189 PROTEIN-TYROSINE PHOSPHATASE.
 FT ACT_SITE 1123 1123 BY SIMILARITY.
 FT DOMAIN 566 573 POLY-PRO.
 FT DOMAIN 635 639 POLY-GLY.
 FT DOMAIN 712 718 POLY-GLU.
 FT SEQUENCE 1189 AA; 135030 MW; 2B85BE5F9C723303 CRC64;

Query Match 57.2%; Score 51.5; DB 1; Length 1189;
 Best Local Similarity 83.3%; Pred. No. 33;
 Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3 RPRPPYLRPRP 14
 11 1111111111
 DB 565 RPRPPY-PRRP 575

RESULT 15
 MK1_PALPR STANDARD; PRT; 15 AA.
 AC P80408;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE METALNIKOVIN I.
 OS Palomena prasina.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomicea;
 OC Pentatomidae; Palomena.
 OX NCBI_TaxID=55431;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
 RT "The inducible antibacterial peptides of the hemipteran insect
 Palomena prasina. Identification of a unique family of proline-rich
 peptides and of a novel insect defensin.";
 RL J. Insect Physiol. 42:81-89(1996).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -1- INDUCTION: BY BACTERIAL INFECTION.
 CC Antibiotic; Insect Immunity.
 CC Antidiotic; Insect Immunity.
 KM SEQUENCE 15 AA; 1838 MW; 21407E663CE46239 CRC64;

Query Match 56.7%; Score 51; DB 1; Length 15;
 Best Local Similarity 72.7%; Pred. No. 0.59;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPPYLRPRRP 15
 1111111111
 DB 3 KPDYLRPRRP 13

RESULT 16
 XG_HUMAN STANDARD; PRT; 180 AA.
 ID XG_HUMAN
 AC P55808;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE XG GLYCOPROTEIN PRECURSOR (PROTEIN PBDX).
 GN XG OR PBDX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=9432149; PubMed=8054981;
 RX Ellis N.A., Ye T.Z., Patton S., German J., Goodfellow P.N.,
 RA Weller P.;
 RT "Cloning of PBDX, an MIC2-related gene that spans the pseudoautosomal
 boundary on chromosome Xp.";
 RT Nat. Genet. 6:394-400(1994).
 RL -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- PTM: O-GLYCOSYLATED (PROBABLE).
 CC -1- POLYMORPHISM: XG IS RESPONSIBLE FOR THE XG BLOOD GROUP SYSTEM.
 CC -1- SIMILARITY: TO PROTEIN MIC2/CD9.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X96421; -; NOT_ANNOTATED_CDS.
 DR MIM; 314700; .
 KM Transmembrane; Glycoprotein; Blood group antigen; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 180 XG GLYCOPROTEIN.
 FT TRANSMEM 143 163 POTENTIAL.
 FT SEQUENCE 180 AA; 19723 MW; DADAA9E6859C4530 CRC64;

Query Match 56.7%; Score 51; DB 1; Length 180;
 Best Local Similarity 53.8%; Pred. No. 6.2;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPRPPYLRPRRP 15
 1111111111
 DB 47 KRPYLRPRRP 59

RESULT 17
 ID C1W4_HUMAN STANDARD; PRT; 393 AA.
 AC 09NYG8;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 4 (TWIK-RELATED ARACHIDONIC ACID-
 DE STIMULATED POTASSIUM CHANNEL PROTEIN) (TRAK).
 GN KCNK4 OR TRAK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX PubMed=11042359;
 RA Chapman C.G., Meadows H.J., Godden R.J., Campbell D.A., Duckworth M.,
 RA Kelsell R.E., Murdoch P.R., Randall A.D., Renne G.I., Gloger I.S.;
 RT "Cloning, localisation and functional expression of a novel human,
 RT cerebellum specific, two pore domain potassium channel.";
 RL Brain Res. Mol. Brain Res. 82:74-83(2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE-Frontal cortex;
 RA Gray A.T.;
 RT *Assignment of KCNK4 encoding the human potassium channel TRAAK to
 CC chromosome 11.*;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: VOLTAGE INSENSITIVE, INSTANTANEOUS, OUTWARDLY RECTIFYING
 POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS REVERSED AT HIGH
 EXTERNAL K+ CONCENTRATIONS (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
 CHANNELS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF248242; AAG31731.1; -
 DR EMBL: AF247042; AAF64062.1; ALT_INIT.
 DR InterPro: IPR001622; -
 DR InterPro: IPR003280; -
 DR PRINTS: PRO1333; 2PORECHANNEL.
 DR Ionic channel; Transmembrane; Ion transport; Potassium transport;
 KW Glycoprotein.
 FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 4 24 POTENTIAL.
 FT DOMAIN 89 113 PORE-FORMING (POTENTIAL).
 FT TRANSSEM 118 138 POTENTIAL.
 FT DOMAIN 140 171 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 172 192 POTENTIAL.
 FT DOMAIN 197 221 PORE-FORMING (POTENTIAL).
 FT TRANSSEM 234 254 POTENTIAL.
 FT DOMAIN 255 393 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 82 82 P -> L (IN REF. 2). (POTENTIAL).
 FT SEQUENCE 328 328
 SQ SEQUENCE 393 AA; 42704 MW; 7F18E53A0A9AD57D CRC64;
 QY 1 RRRPRPYLP-RRPR 15
 Db 368 RRRPNPRKPRVRRPGR 383
 Query Match 56.1%; Score 50.5; DB 1; Length 393;
 Best Local Similarity 68.8%; Pred. No. 15;
 Matches 11; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 RESULT 18
 APID_BOMPA
 AC P81464; STANDARD; PRT; 17 AA.
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE APIDAEICIN.
 OS Bombus pascuorum.
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
 CC Apoidea; Apidae; Bombus.
 CC NCBI_TaxID=65598;
 CX [1]
 FN SEQUENCE.
 FP TISSUE-Hemolymph;
 FC MEDLINE=97362903; PubMed=9219367;
 EX Rees J.A., Montal M., Buret P.;
 RA "Novel antibacterial peptides isolated from a European bumblebee,
 RT Bombus pascuorum (Hymenoptera, Apoidea).";
 RT Insect Biochem. Mol. Biol. 27:413-422(1997).
 RN

CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -1- INDUCTION: BY BACTERIAL INFECTION.
 CC Insect Immunol; Antibiotic; Hemolymph.
 KW SEQUENCE 17 AA; 1963 MW; CD1DD02C8BC23D1 CRC64;
 SQ
 QY 5 RPYLPRRPP 15
 Db 3 RPYLPRRPP 13
 Query Match 55.6%; Score 50; DB 1; Length 17;
 Best Local Similarity 72.7%; Pred. No. 0.87;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 RESULT 19
 ACRO_PIG
 ID ACRO_PIG STANDARD; PRT; 415 AA.
 AC P08001; P08000;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ACROSIN PRECURSOR (EC 3.4.21.10) (53 KDA FUCOSE-BINDING PROTEIN).
 GN ACR.
 OS Sus scrofa (Pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 CC NCBI_TaxID=9823;
 CX [1]
 RN MEDLINE=89308595; PubMed=2745422;
 RX Baba T., Kashiwabara S.I., Watanabe K., Itoh H., Michikawa Y.,
 RA Kimura K., Takada M., Fukamizu A., Arai Y.;
 RT "Activation and maturation mechanisms of boar acrosin zymogen based
 RT on the deduced primary structure.";
 RL J. Biol. Chem. 264:11920-11927(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE=89325501; PubMed=2502391;
 RA Adham I.M., Maier W.-M., Hoyer-Fender S., Tsaousidou S., Engel W.,
 RA Klemm U.;
 RT "Molecular cloning of preproacrosin and analysis of its expression
 RT pattern in spermatogenesis.";
 RL Eur. J. Biochem. 182:563-568(1989).
 RN [3]
 RP SEQUENCE OF 17-39.
 RC TISSUE-Sperm;
 RX MEDLINE=84261484; PubMed=6378631;
 RA Fock-Nuzel R., Lottspeich F., Henschen A., Mueller-Esterl W.;
 RT "Boar acrosin is a two-chain molecule. Isolation and primary
 RT structure of the light chain; homology with the pro-part of other
 RT serine proteinases.";
 RL Eur. J. Biochem. 141:441-446(1984).
 RN [4]
 RP SEQUENCE OF 25-91.
 RC TISSUE-Sperm;
 RX MEDLINE=8115822; PubMed=7007202;
 RA Fock-Nuzel R., Lottspeich F., Henschen A., Mueller-Esterl W.,
 RA Fritz H.;
 RT "N-Terminal amino acid sequence of boar sperm acrosin. Homology with
 RT other serine proteinases.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 361:1823-1828(1980).
 RN [5]
 RP SEQUENCE OF 17-32 AND 40-53.
 RC TISSUE-Sperm;
 RX MEDLINE=88083633; PubMed=3480243;
 RA Toepfer-Petersen E., Henschen A.;
 RT "Acrosin shows zona and fucose binding, novel properties for a serine
 RT proteinase.";
 RL FEBS Lett. 226:38-42(1987).
 RN [6]

RP SEQUENCE OF 17-40.
 RA MEDLINE-90253655; PubMed-2111146;
 RX Cechova D., Toepfer-Petersen E., Zuckner A., Jonakova V.;
 RT "Is spermogen a modified proacrosin? Isolation, purification, and
 partial characterization of low-molecular-mass boar proacrosin.";
 RL Biol. Chem. Hoppe-Seyler 371:317-323(1990).
 RN [7].
 RP DISULFIDE BONDS, CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
 RA MEDLINE-91085546; PubMed-2261983;
 RX Toepfer-Petersen E., Calvete J.J., Schaefer W., Henschen A.;
 RT "Complete localization of the disulfide bridges and glycosylation
 sites in boar sperm acrosin.";
 RL FEBS Lett. 275:139-142(1990).
 CC -1- FUNCTION: ACROSIN IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.
 CC IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY. IT
 CC IS SYNTHESIZED IN A ZYMOGEN FORM, PROACROSIN AND STORED IN THE
 CC ACROSOME.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL
 CC CLEAVAGE ARG-XAA->> LYS-LYS->> LYS-XAA.
 CC -1- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO
 CC DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J04950; AAA3131.1; -;
 DR EMBL: X14844; CAA32948.1; -;
 DR PIR: A34170; A34170.
 DR PIR: S02428; S02428.
 DR PIR: S04940; S04940.
 DR PIR: S08994; S08994.
 DR PIR: S12968; S12968.
 DR MEROPS: S01.223; -;
 DR InterPro: IPR001254; -;
 DR InterPro: IPR001314; -;
 DR Pfam: PF00089; Trypsin. 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 415 ACROSIN.
 FT CHAIN 17 39 ACROSIN LIGHT CHAIN.
 FT CHAIN 40 338 ACROSIN HEAVY CHAIN.
 FT PROPEP 339 415 PRO-RICH.
 FT DISULFID 22 152 INTERCHAIN.
 FT DISULFID 26 160 INTERCHAIN.
 FT DISULFID 71 87
 FT DISULFID 175 244
 FT DISULFID 207 223
 FT DISULFID 234 264
 FT CARBOHYD 19 19
 FT CARBOHYD 208 208
 FT ACT_SITE 86 86 N-LINKED (GLCNAC. . .).
 FT ACT_SITE 140 140 CHANGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 238 238 CHANGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 8 8 MISSING (IN REF. 2).
 FT CONFLICT 211 211 R->O (IN REF. 2).
 FT CONFLICT 217 218 IR->VT (IN REF. 2).
 FT CONFLICT 347 347 P->A (IN REF. 2).
 FT CONFLICT 389 389 MISSING (IN REF. 2).
 FT CONFLICT 399 402 RSYV->KELL (IN REF. 2).
 FT CONFLICT 403 415 MISSING (IN REF. 2).
 SO SEQUENCE 415 AA; 45387 MW; 5AD11900B7E95772 CRC64;

Query Match 55.6%; Score 50; DB 1; Length 415;
 Best Local Similarity 69.2%; Pred. No. 18;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 3 RRRPPPLPRPP 15
 |||||
 Db 338 RRRPPAPPPPP 350
 RESULT 20
 SP49_HUMAN STANDARD; PRT; 424 AA.
 ID SP49_HUMAN
 AC Q15427;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SPICEOSOME ASSOCIATED PROTEIN 49 (SAP 49) (SF3B53).
 GN SF3B4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95047348; PubMed-7958871;
 RA Champion-Arnaud P., Reed R.;
 RT "The prespliceosome components SAP 49 and SAP 145 interact in a
 RT complex implicated in tethering U2 snRNP to the branch site.";
 RL Genes Dev. 8:1974-1983(1994).
 CC -1- FUNCTION: SUBUNIT OF THE SF3B COMPLEX REQUIRED FOR 'A' COMPLEX
 CC ASSEMBLY FORMED BY THE STABLE BINDING OF U2 snRNP TO THE
 CC BRANCHPOINT SEQUENCE (BPS) IN PRE-MRNA. SEQUENCE INDEPENDENT
 CC BINDING OF SF3A/SF3B COMPLEX UPSTREAM OF THE BRANCH SITE IS
 CC ESSENTIAL. IT MAY ANCHOR U2 snRNP TO THE PRE-MRNA. SAP 49 HAS BEEN
 CC FOUND IN COMPLEX 'B' AND 'C' AS WELL.
 CC -1- SUBUNIT: SF3B IS COMPOSED OF SAPS 49, 130, 145, AND 155 (SF3B53,
 CC SF3B120, SF3B150 AND SF3B169 RESPECTIVELY). SAP 49 INTERACTS
 CC DIRECTLY WITH SAP 145.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS COB11.5 AND, PARTIAL, TO
 CC YEAST HSH49.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L35013; AAA60300.1; -;
 DR HSP: P19339; 2SXU.
 DR InterPro: IPR00504; -;
 DR Pfam: PF00076; Rtm; 2.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW mRNA processing; RNA-binding; Nuclear protein; Repeat.
 FT DOMAIN 13 91 RNA-BINDING (RRM) 1.
 FT DOMAIN 100 179 RNA-BINDING (RRM) 2.
 FT DOMAIN 215 218 POLY-PRO.
 FT DOMAIN 262 268 POLY-PRO.
 SO SEQUENCE 424 AA; 44385 MW; 212472A25D3FF002 CRC64;

Query Match 55.6%; Score 50; DB 1; Length 424;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 RRRPPPLPRPP 15
 :|||
 Db 397 QRGPLPPPRPP 411

```

RESULT 21
APG_BRANA STANDARD: PRT: 449 AA.
AC 027905:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ANTER-SPECIFIC PROLINE-RICH PROTEIN APG (PROTEIN CEX) (FRAGMENT).
GN APG OR CEX.
CS Brassica napus (Rape).
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
CC Brassicales; Brassicaceae; Brassica.
CX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94004980; PubMed=8401599;
RA Roberts M.R., Foster G.D., Blundell R.P., Robinson S.W., Kumar A.,
RA Draper J., Scott R.;
RT "gametophytic and sporophytic expression of an anther-specific
RT Arabidopsis thaliana gene.";
RL Plant J. 3:111-120(1993).
CC -1- TISSUE SPECIFICITY: FOUND IN ANTHER, ONLY IN MALE FERTILE PLANTS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN MALE GAMETOGENESIS, DURING
CC MICROSPORE DEVELOPMENT. HIGHER EXPRESSION IS FOUND DURING
CC MICROSPORE MITOSIS WITH A DRAMATIC DECLINE DURING POLLEN
CC MATURATION.
CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60376; CAA42924.1; -
DR PIR: S16748; S16748.
DR InterPro: IPR001087; -
DR Pfam: PF00657; Lipase_GDSL; 1.
DR PROSITE: PS01098; LIPASE_GDSL_SER; 1.
FT NON_TER 1 1
FT ACT_SITE 132 132 BY SIMILARITY.
FT ACT_SITE 428 428 POTENTIAL.
FT ACT_SITE 449 AA: 48779 MW; 9EPB6A3AB28EA15 CRC64;
SQ
Query Match 55.6%; Score 50; DB 1; Length 449;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 3 RPPYLP RPPPP 15
DB 5 QPKPPKQPKRP 17
RESULT 22
ABPP_RIPCL STANDARD: PRT: 678 AA.
AC 027905:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE ANTIBACTERIAL PEPTIDE POLYPROTEIN PRECURSOR.
OS Riptortus clavatus.
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
CC Heteroptera; Panheteroptera; Pentatomomorpha; Coreoidea; Alydidae;
CC Riptortus.
CX NCBI_TaxID=41704;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Fat body; PubMed=8688805;
RX MEDLINE=96261233;
RA Miura K., Ueno S., Kamiya K., Kobayashi J., Matsuo H., Ando K.,
RA Chizel Y.;
RT "Cloning of mRNA sequences for two antibacterial peptides in a
RT hemipteran insect, Riptortus clavatus.";
RL Zool. Sci. 13:111-117(1996).
CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY IN VITRO.
CC -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -1- SIMILARITY: TO PYRROCORICIN, DROSOCIN AND APTADACIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D49415; BAA08395.1; -
KW Polyprotein; Insect immunity; 14 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 1 678
FT REPEAT 1 67 1-1.
FT REPEAT 68 114 1-2.
FT REPEAT 115 161 1-3.
FT REPEAT 162 208 1-4.
FT REPEAT 209 255 1-5.
FT REPEAT 256 302 1-6.
FT REPEAT 303 349 1-7.
FT REPEAT 350 396 1-8.
FT REPEAT 397 443 1-8.
FT REPEAT 444 490 1-10.
FT REPEAT 491 537 1-11.
FT REPEAT 538 584 1-12.
FT REPEAT 585 631 1-13.
FT REPEAT 632 678 1-14.
FT CARBOHYD 32 32 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 83 83 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 224 224 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 506 506 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 600 600 O-LINKED (GALNAC. . .) (POTENTIAL).
FT CARBOHYD 647 647 O-LINKED (GALNAC. . .) (POTENTIAL).
SQ SEQUENCE 678 AA: 76367 MW; 2939BA68692D2444 CRC64;
Query Match 55.6%; Score 50; DB 1; Length 678;
Best Local Similarity 72.7%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 5 RPPYLP RPPPP 15
DB 24 KPPYLP RPPPP 34
RESULT 23
GDA6_WHEAT STANDARD: PRT: 296 AA.
AC P04726:
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA/BETA-GLIADIN CLONE PM1215 PRECURSOR (PROMALIN).
OS Triticum aestivum (Wheat).
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

CC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae;
 CC Triticum.
 CC NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=65242077; PubMed=3839304;
 RA Sumner-Smith M., Ratajski J.A., Sugiyama T., Stoll M., Soell D.;
 RT "Conservation and variability of wheat alpha/beta-gliadin genes";
 RL Nucleic Acids Res. 13:3905-3916(1985).
 CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
 CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
 CC HOMOLOG CLASSES: SEQUENCE DIVERGENCE BETWEEN THE CLASSES IS DUE
 CC TO SINGLE BASE SUBSTITUTIONS AND TO DUPLICATIONS OR DELETIONS
 CC WITHIN OR NEAR DIRECT REPEATS. THERE ARE MORE THAN A 100 COPIES OF
 CC THE GENE FOR ALPHA/BETA-GLIADIN PER HAPLOID GENOME.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X02538; CA26383.1; -;
 CC EMBL: K03074; AAA34277.1; -;
 CC PIR: S07361; S07361.
 CC InterPro: IPR001376; -;
 CC InterPro: IPR001768; -;
 CC InterPro: IPR001954; -;
 CC Pfam: PF00234; ttyp_alpha.amyl.1.
 CC PRINTS: PR00208; GLIADIN.LUTEN.
 CC PRINTS: PR00209; GLIADIN.
 CC Seed storage protein; Repeat; Signal; Multigene family.
 KW SIGNAL
 FT CHAIN 1 296 ALPHA/BETA-GLIADIN CLONE PW1215.
 FT SEQUENCE 296 AA: 33941 MW: 49590AD40F135 CRC64;
 SQ
 QY 3 RRRPPYLRPPRP 15
 DB 85 QRPPLPQPPQPP 97
 Query Match 54.4%; Score 49; DB 1; Length 296;
 Best Local Similarity 61.5%; Pred. No. 17;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RA Tsipouras P., Ramirez F., Hollister D.;
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to
 RT two different fibrillin genes";
 RL Nature 352:330-334(1991).
 RN [3]
 RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
 RX MEDLINE=96083599; PubMed=7493032;
 RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
 RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
 RT congenital contractural arachnodactyly";
 RL Nat. Genet. 11:456-458(1995).
 CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
 CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
 CC -1- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
 CC ARACHNOACTYL (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
 CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
 CC AORTA AND THE EYES.
 CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
 CC EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U03272; AAA18950.1; -;
 CC EMBL: X62009; -; NOT_ANNOTATED_CDS.
 CC PIR: S17063; S17063.
 CC PIR: S31101; S31101.
 CC HSP: P35555; IEMO.
 CC MIM: 121050; -;
 CC InterPro: IPR000152; -;
 CC InterPro: IPR000561; -;
 CC InterPro: IPR001438; -;
 CC InterPro: IPR001881; -;
 CC InterPro: IPR002212; -;
 CC Pfam: PF00008; EGF; 46.
 CC Pfam: PF00683; TB; 9.
 CC PRINTS: PR00010; EGF.BLOOD.
 CC PROSITE: PS00010; ASX_HYDROXYL; 43.
 CC PROSITE: PS00022; EGF_1; 2.
 CC PROSITE: PS01186; EGF_2; 37.
 CC PROSITE: PS01187; EGF_CA; 43.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
 FT SIGNAL 1 28
 FT CHAIN 29 2911
 FT DOMAIN 111 142
 FT DOMAIN 145 176
 FT DOMAIN 176 207
 FT DOMAIN 275 316
 FT DOMAIN 317 358
 FT DOMAIN 359 425
 FT REPEAT 493 533
 FT DOMAIN 534 573
 FT DOMAIN 574 615
 FT DOMAIN 616 656
 FT DOMAIN 657 697
 FT REPEAT 698 766
 FT DOMAIN 767 808
 FT DOMAIN 809 850
 FT DOMAIN 851 890
 FT DOMAIN 954 995
 FT REPEAT 996 1071
 FT DOMAIN 1072 1113
 FT DOMAIN 1114 1156
 FT DOMAIN 1157 1198
 FT DOMAIN 1199 1240
 EGF-LIKE 1, NON-CALCIUM BINDING.
 EGF-LIKE 2, NON-CALCIUM BINDING.
 EGF-LIKE 3, NON-CALCIUM BINDING.
 EGF-LIKE 4, CALCIUM-BINDING.
 EGF-LIKE 5, CALCIUM-BINDING.
 TGF-1.
 EGF-LIKE 6, NON-CALCIUM BINDING.
 EGF-LIKE 7, CALCIUM-BINDING.
 EGF-LIKE 8, CALCIUM-BINDING.
 EGF-LIKE 9, CALCIUM-BINDING.
 EGF-LIKE 10, CALCIUM-BINDING.
 TGF-2.
 EGF-LIKE 11, CALCIUM-BINDING.
 EGF-LIKE 12, CALCIUM-BINDING.
 EGF-LIKE 13, CALCIUM-BINDING.
 EGF-LIKE 14, CALCIUM-BINDING.
 TGF-3.
 EGF-LIKE 15, CALCIUM-BINDING.
 EGF-LIKE 16, CALCIUM-BINDING.
 EGF-LIKE 17, CALCIUM-BINDING.
 EGF-LIKE 18, CALCIUM-BINDING.


```

FT DOMAIN 1241 1281 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1282 1323 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1324 1365 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1366 1406 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1407 1447 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1448 1489 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1490 1530 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1531 1571 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1572 1648 TGFBR 4.
FT REPEAT 1649 1690 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1691 1732 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1733 1806 TGFBR 5.
FT REPEAT 1807 1848 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1891 1932 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1933 1971 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1972 2014 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2015 2054 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2055 2096 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2097 2169 TGFBR 6.
FT DOMAIN 2170 2211 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2212 2251 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2252 2292 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2293 2336 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2337 2378 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2379 2447 TGFBR 7.
FT DOMAIN 2448 2489 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2490 2530 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2531 2569 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2570 2612 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2613 2652 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2653 2693 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2694 2733 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
FT DISULFID 132 141 BY SIMILARITY.
FT DISULFID 149 159 BY SIMILARITY.
FT DISULFID 153 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 184 195 BY SIMILARITY.
FT DISULFID 197 206 BY SIMILARITY.
FT DISULFID 206 2291 BY SIMILARITY.
FT DISULFID 279 291 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 302 315 BY SIMILARITY.
FT DISULFID 321 333 BY SIMILARITY.
FT DISULFID 328 342 BY SIMILARITY.
FT DISULFID 344 357 BY SIMILARITY.
FT DISULFID 497 509 BY SIMILARITY.
FT DISULFID 504 518 BY SIMILARITY.
FT DISULFID 520 532 BY SIMILARITY.
FT DISULFID 538 548 BY SIMILARITY.
FT DISULFID 543 557 BY SIMILARITY.
FT DISULFID 559 572 BY SIMILARITY.
FT DISULFID 578 590 BY SIMILARITY.
FT DISULFID 585 599 BY SIMILARITY.
FT DISULFID 601 614 BY SIMILARITY.
FT DISULFID 620 631 BY SIMILARITY.
FT DISULFID 626 640 BY SIMILARITY.
FT DISULFID 642 655 BY SIMILARITY.
FT DISULFID 661 672 BY SIMILARITY.
FT DISULFID 667 681 BY SIMILARITY.
FT DISULFID 683 696 BY SIMILARITY.
FT DISULFID 771 783 BY SIMILARITY.
FT DISULFID 778 792 BY SIMILARITY.
FT DISULFID 794 807 BY SIMILARITY.
FT DISULFID 813 825 BY SIMILARITY.
FT DISULFID 820 834 BY SIMILARITY.
FT DISULFID 836 849 BY SIMILARITY.
FT DISULFID 855 865 BY SIMILARITY.
FT DISULFID 860 874 BY SIMILARITY.
FT DISULFID 876 889 BY SIMILARITY.
FT DISULFID 958 970 BY SIMILARITY.

```

```

FT DISULFID 965 979 BY SIMILARITY.
FT DISULFID 981 994 BY SIMILARITY.
FT DISULFID 1076 1088 BY SIMILARITY.
FT DISULFID 1083 1097 BY SIMILARITY.
FT DISULFID 1099 1112 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1125 1139 BY SIMILARITY.
FT DISULFID 1141 1155 BY SIMILARITY.
FT DISULFID 1161 1173 BY SIMILARITY.
FT DISULFID 1168 1182 BY SIMILARITY.
FT DISULFID 1184 1197 BY SIMILARITY.
FT DISULFID 1203 1215 BY SIMILARITY.
FT DISULFID 1210 1224 BY SIMILARITY.
FT DISULFID 1226 1239 BY SIMILARITY.
FT DISULFID 1245 1256 BY SIMILARITY.
FT DISULFID 1252 1265 BY SIMILARITY.
FT DISULFID 1267 1280 BY SIMILARITY.
FT DISULFID 1286 1298 BY SIMILARITY.
FT DISULFID 1293 1307 BY SIMILARITY.
FT DISULFID 1309 1322 BY SIMILARITY.
FT DISULFID 1328 1340 BY SIMILARITY.
FT DISULFID 1335 1349 BY SIMILARITY.
FT DISULFID 1351 1364 BY SIMILARITY.
FT DISULFID 1370 1383 BY SIMILARITY.
FT DISULFID 1377 1392 BY SIMILARITY.
FT DISULFID 1394 1405 BY SIMILARITY.
FT DISULFID 1411 1424 BY SIMILARITY.
FT DISULFID 1418 1433 BY SIMILARITY.
FT DISULFID 1435 1446 BY SIMILARITY.
FT DISULFID 1452 1464 BY SIMILARITY.
FT DISULFID 1459 1473 BY SIMILARITY.

```

```

Query Match 54.48; Score 49; DB 1; Length 2911;
Best Local Similarity 61.5%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 3 RPPPPYLP RPP 15
Db 31 OPPPPPP RPP 43

```

```

RESULT 25
BAT2_HUMAN STANDARD; PRT; 2142 AA.
ID BAT2_HUMAN
AC P46634;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE LARGE PROLINE-RICH PROTEIN BAT2 (HLA-B-ASSOCIATED TRANSCRIPT 2).
GN BAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RA MEDLINE=90192810; PubMed=2156268;
RA Banerji J., Sands J., Strominger J.L., Spies T.;
RT "A gene pair from the human major histocompatibility complex encodes
RT large proline-rich proteins with multiple repeated motifs and a
RT single ubiquitin-like domain."
RT Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
RL [2]
RN SEQUENCE OF 1-1860 FROM N.A.
RP MEDLINE=93272029; PubMed=849947;
RA Irls F.J.M., Bougueleret L., Prieur S., Caterina D., Primas G.,
RA Petrot V., Jurka J., Rodriguez-Tome P., Claverie J.-M., Dausset J.,
RA Cohen D.;
RT "Dense Alu clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment."
RL Nat. Genet. 3:137-145(1993).
CC -1- FUNCTION: UNKNOWN.

```

CC -1- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKEMIC ORIGIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M33509; AAA35585.1; -
 DR EMBL; M33518; AAA35586.1; -
 DR EMBL; M33512; AAA35586.1; JOINED.
 DR EMBL; Z15025; CAA78744.1; -
 DR PIR; B35098; B35098.
 DR PIR; S36152; S36152.
 DR MIM; 142580; -
 KW Repeat.
 FT DOMAIN 519 524 POLY-PRO.
 FT DOMAIN 636 657 GLN-RICH.
 FT DOMAIN 684 688 POLY-PRO.
 FT DOMAIN 699 704 POLY-PRO.
 FT DOMAIN 814 821 POLY-PRO.
 FT DOMAIN 1340 1345 POLY-GLY.
 FT DOMAIN 1398 1403 POLY-GLY.
 FT DOMAIN 1436 1442 POLY-PRO.
 FT DOMAIN 1982 1991 POLY-PRO.
 FT DOMAIN 41 1795 4 X 57 AA TYPE A REPEATS.
 FT REPEAT 41 95 1-1.
 FT REPEAT 98 154 1-2.
 FT REPEAT 281 337 1-3.
 FT REPEAT 1740 1795 1-4.
 FT DOMAIN 337 549 2 X TYPE B REPEATS.
 FT REPEAT 337 418 2-1.
 FT REPEAT 476 549 2-2.
 FT DOMAIN 1899 2089 3 X 50 AA TYPE C REPEATS.
 FT REPEAT 1899 1948 3-1.
 FT REPEAT 1965 2014 3-2.
 FT REPEAT 2040 2089 3-3.
 FT CONFLICT 57 57 R -> A (IN REF. 2).
 FT CONFLICT 109 109 Q -> S (IN REF. 2).
 FT CONFLICT 414 414 P -> PPHGPGAGNMGPP (IN REF. 2).
 FT CONFLICT 532 532 T -> K (IN REF. 2).
 FT CONFLICT 682 682 Q -> D (IN REF. 2).
 FT CONFLICT 730 730 E -> D (IN REF. 2).
 FT CONFLICT 750 750 L -> R (IN REF. 2).
 FT CONFLICT 834 834 A -> T (IN REF. 2).
 FT CONFLICT 1035 1035 G -> A (IN REF. 2).
 FT CONFLICT 1068 1068 M -> L (IN REF. 2).
 FT CONFLICT 1285 1285 P -> R (IN REF. 2).
 FT CONFLICT 1400 1400 G -> A (IN REF. 2).
 FT CONFLICT 1611 1611 T -> S (IN REF. 2).
 FT CONFLICT 1729 1729 G -> A (IN REF. 2).
 SO SEQUENCE 2142 AA; 227840 MW; 32DDF16B9B52420A CRC64;

Query Match 53.9%; Score 48.5; DB 1; Length 2142;
 Best Local Similarity 61.1%; Pred. No. 1.3e+02;
 Matches 11; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 QY 1 RRRP---RPPYLRPPRP 15
 Db 1424 RSRPERRPPGLRPPPP 1441

RESULT 26
 FASL.HUMAN STANDARD; PRT; 281 AA.
 AC F48023;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE FAS ANTIGEN LIGAND (APOPTOSIS ANTIGEN LIGAND) (APTL).

GN TNFSF6 OR APTLI1 OR FASL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95105731; PubMed=7528780;
 RA Alderson M.;
 RT "Fas ligand mediates activation-induced cell death in human T
 lymphocytes.";
 RL J. Exp. Med. 181:71-77(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127560; PubMed=7826947;
 RA Takahashi T.; Tanaka M.; Inazawa J.; Abe T.; Suda T.; Nagata S.;
 RT "Human Fas ligand: gene structure, chromosomal location and species
 specificity.";
 RL Int. Immunol. 6:1567-1574(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Schaeuble C.E.; Poehlmann R.; Philippson P.; Eitel H.;
 RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9507350; PubMed=7980502;
 RA Mita E.; Hayashi N.; Ito S.; Takehara T.; Hijioka T.; Kasahara A.;
 RT "Role of Fas ligand in apoptosis induced by hepatitis C virus
 infection.";
 RL Biochem. Biophys. Res. Commun. 204:468-474(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX Wilkinson J.;
 RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC Tissue-Blood;
 RA Matsumura M.; Nakanishi Y.; Ohba Y.;
 RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
 CC TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
 CC CYTOTOXIC T CELL-MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
 CC FAS-ANTIGEN-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
 CC PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
 CC T CELLS, OR BOTH.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
 CC INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
 CC SURFACE.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X89102; CAA61474.1; -
 DR EMBL; U08137; AAC50071.1; -
 DR EMBL; U11821; AAC50124.1; -
 DR EMBL; D38122; BAA07320.1; -
 DR EMBL; Z96050; CAB09424.1; -
 DR EMBL; AB013303; BAA32542.1; -
 DR HSSP; P01375; 2FUN.
 DR MIM; 134638; -
 DR InterPro: IPR000478; -
 DR Pfam: PF00229; TNF_1; 1.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Apoptosis.

```

FT DOMAIN 1 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 102 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 103 281 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 4 70 PRO-RICH.
FT DOMAIN 45 65 POLY-PRO.
FT DISULFID 202 233 BY SIMILARITY.
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 281 AA; 31485 MW; ABA6EB358246E9BB CRC64;

Query Match 53.3%; Score 48; DB 1; Length 281;
Best Local Similarity 60.0%; Pred. NO. 21;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRRPPPYLPPRPP 15
    ||| || | ||
    42 QRRPPPPPPPLPP 56

RESULT 27
VGLD_PRIVRI STANDARD; PRT; 402 AA.
ID VGLD_PRIVRI STANDARD; PRT; 402 AA.
AC P07645;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLYCOPROTEIN GP50.
OC Pseudorabies virus (strain Rice) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OK NCBI_TaxID=10350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66281819; PubMed=3016293;
RX Petrovskis E.A., Timmins J.G., Armentrout M.A., Marcholi C.C.,
RA Vancey R.J. Jr., Post L.E.;
RT "DNA sequence of the gene for pseudorabies virus gp50, a glycoprotein
  without N-linked glycosylation."
RL J. Virol. 59:216-223(1986).
- - MISCELLANEOUS: GLYCOPROTEIN GP50 DOES NOT CONTAIN N-LINKED
  CARBOHYDRATE, AS PREDICTED FROM ITS SEQUENCE.
- - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN D FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M14001; AAC35203.1; -
CC PIR; A27788; VGBE50.
CC InterPro: IPR002896; -
CC Pfam: PF01537; Herpes_glycop_D; 1.
CC Glycoprotein
KW SEQUENCE 402 AA; 44501 MW; B876330595871E8 CRC64;

Query Match 53.3%; Score 48; DB 1; Length 402;
Best Local Similarity 75.0%; Pred. NO. 30;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RRRPPPYLPPRPP 14
    ||| ||| |||
    268 RRRPPPPPPRPP 279

RESULT 28
CP84_ARATH STANDARD; PRT; 520 AA.
ID CP84_ARATH STANDARD; PRT; 520 AA.

```

```

AC Q42600;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOCHROME P450 84A1 (FERULATE-5-HYDROXYLASE) (EC 1.14.-.-) (FSH).
GN CYP84A1 OR F4H1 OR A74G36220 OR F23E13.110.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=96293440; PubMed=8692910;
RA Meyer K., Cusumano J.C., Somerville C.R., Chapple C.C.S.;
RT "Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new
  family of cytochrome P450-dependent monooxygenases."
RL Proc. Natl. Acad. Sci. U.S.A. 93:6869-6874(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Griwell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaler B., Maché R., Mueller M.,
RA Kreis M., Delseny M., Pilgomench P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
RA Breken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moellmann P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lambirth S., Van den Daele H.,
RA De Keyser A., Buysheert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Petzelt A., Rajadream M.-A., Lyne M., Benes V., Rechmann S.,
RA Botkova D., Bloeker H., Scharfe M., Grimm M., Loehert T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argiriou A., Vitale D., Liguori R., Pirvandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Raon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Frans P., Biele C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedha N., Gnoj L., Schutz K., Hang E., Spiegel L.,
RA Sehnok M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harnon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Gaisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vall D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez S., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodi M., Johnson A.,
RA Chen E., Maria M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
  thaliana."
RL Nature 402:769-777(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERRECTA;
RX MEDLINE=99097044; PubMed=9880351;

```

```

RA Ruegger M., Meyer K., Cusumano J.C., Chapple C.;
RT "The regulation of ferulate-5-hydroxylase expression in Arabidopsis in
CC the context of sinapate ester biosynthesis.";
RL Plant Physiol. 119:101-110(1999).
CC -1- PATHWAY: GENERAL PHENYLPROANOIC PATHWAY
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U38416; AAC49389.1; -.
DR EMBL: AL022141; CAA18128.1; -.
DR EMBL: AF068574; AAD11580.1; -.
DR EMBL: AL161389; CAB80293.1; -.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Membrane; Heme.
KM BINDING 458 458 HEME (BY SIMILARITY).
SQ SEQUENCE 520 AA; 58720 MW; E812779AF5BF01BC CRC64;

QY 2 RRRPPYLRPR 13
   || || || || ||
DB 34 RRRRPPYPRGR 45

Query Match          53.3%; Score 48; DB 1; Length 520;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 29
TOP3_CAEEL          STANDARD;      PRT;      759 AA.
AC 061660;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA TOPOISOMERASE III (EC 5.99.1.2).
GN TOP3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN RP
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Kim Y.-C., Koo H.-S.;
RT "cDNA cloning and overexpression of Caenorhabditis elegans DNA
RT topoisomerase III.";
RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE OF SINGLE-STRANDED
CC DNA, FOLLOWED BY PASSAGE AND REJOINING.
CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF057032; AAC13567.1; -.
DR InterPro: IPR000380; -.
DR InterPro: IPR002936; -.

```

```

DR Pfam: PF01131; Topoisom_bac; 1.
DR Pfam: PF01751; Toprim; 1.
DR PRINTS: PR00417; PRTPISMRASPI.
DR PROSITE: PS00396; TOPOISOMERASE_I_PROK; 1.
KM Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 334 334 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 759 AA; 85438 MW; 3D862412D72946BD CRC64;

QY 4 RPPYLRPRRP 15
   || || || || ||
DB 646 RPPAPKPRGRPP 657

Query Match          53.3%; Score 48; DB 1; Length 759;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 30
EXTN_SORBI          STANDARD;      PRT;      283 AA.
ID EXTN_SORBI
AC P24152;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
GN HRCF.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN RP
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RC MEDLINE=91370882; PubMed=1893107;
RA Raz R., Crelin C., Puigdemonech P., Martinez-Izquierdo J.A.;
RT "The sequence of a hydroxyproline-rich glycoprotein gene from Sorghum
RT vulgare.";
RL Plant Mol. Biol. 16:365-367(1991).
CC -1- FUNCTION: STRUCTURAL COMPONENT IN PRIMARY CELL WALL.
CC -1- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
CC GLYCOSYLATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56010; CAA39485.1; -.
DR PIR: S14449; S14449.
DR InterPro: IPR002965; -.
DR PRINTS: PR01217; PRICHEXTENS.
KM Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
KM Hydroxylation.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 283 EXTENSIN.
SQ SEQUENCE 283 AA; 29593 MW; 8D7FCD0DA8ED2D90 CRC64;

QY 3 RPPRP-VLRPRRP 15
   || || || || || ||
DB 162 RPPDPVTPNPKPP 175

Query Match          52.8%; Score 47.5; DB 1; Length 283;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

```

```

RESULT 31
DBPA.HUMAN STANDARD: PRT: 372 AA.
AC P16969; Q14121;
DT 01-AUG-1990 (Rel. 15, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA-BINDING PROTEIN A (COLD SHOCK DOMAIN PROTEIN A) (SINGLE-STRAND
GN DNA BINDING PROTEIN NF-GMB).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (FORM 3).
RA MEDLINE=89211987; PubMed=2977358;
RA Sakuta H., Maekawa T., Imamoto F., Yasuda K., Ishii S.;
RT "Two human genes isolated by a novel method encode DNA-binding
RT proteins containing a common region of homology.";
RL Gene 73:499-507(1988).
RN [2]
RP SEQUENCE FROM N.A. (FORMS 2 AND 3).
RA TISSUE=Placenta;
RX MEDLINE=95354705; PubMed=7628487;
RX Kudo S., Mattei M.-G., Fukuda M.;
RT "Characterization of the gene for dbpa, a family member of the
RT nucleic-acid-binding proteins repress transcription from the GM-CSF
RT bur. J. Biochem. 231:72-82(1995).
RN [3]
RP SEQUENCE FROM N.A. (FORM 1).
RX MEDLINE=96279731; PubMed=8710501;
RX Coles L.S., Diamond P., Occhiodoro F., Vadas M.A., Shannon M.F.;
RT "Cold shock domain proteins repress transcription from the GM-CSF
RT promoter.";
RL Nucleic Acids Res. 24:2311-2317(1996).
CC -1- FUNCTION: BINDS TO THE GM-CSF PROMOTER. SEEMS TO ACT AS A
CC REPRESSOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2
CC AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE AND HEART.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M24069; AAA35749.1; ALT.INCT.
DR EMBL: L29071; AAA79243.1; -.
DR EMBL: L29064; AAA79243.1; JOINED.
DR EMBL: L29065; AAA79243.1; JOINED.
DR EMBL: L29066; AAA79243.1; JOINED.
DR EMBL: L29067; AAA79243.1; JOINED.
DR EMBL: L29068; AAA79243.1; JOINED.
DR EMBL: L29069; AAA79243.1; JOINED.
DR EMBL: L29070; AAA79243.1; JOINED.
DR EMBL: X95325; CAA64631.1; -.
DR PIR: PS0014; PS0014.
DR HSSP: P15277; IMTC.
DR TRANSFAC: T00185; -.
DR MIM: 603437; -.
DR InterPro: IPR002059; -.
DR Pfam: PF00313; CSD. 1.
DR PRINTS: PRO0050; COLD_SHOCK.
DR PROSITE: PS00352; COLD_SHOCK. 1.
DR Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Alternative splicing.
FT DOMAIN 93 157 CSD.

```

```

FT VARSPIC 192 260 MISSING (IN ISOFORM 2).
FT VARSPIC 340 372 REPNASDGDKEAKAGEAPTEENAPPTQSSAE -> PSS
FT CONFLICT 75 75 A -> T (IN REF. 3).
SQ SEQUENCE 372 AA; 40060 MW; A52402AC3F6385CC CRC64;

Query Match 52.8%; Score 47.5; DB 1; Length 372;
Best Local Similarity 36.7%; Pred. No. 32;
Matches 11; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

DY 1 RRRPRPYLP-----RRPP 15
ID ||||| |
DB 336 RRRPRPNPSSDGDKEAKAGEAPTEENAPPP 365

RESULT 32
SYJ2.HUMAN STANDARD: PRT: 1443 AA.
AC 015056;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SYNAPTOJANIN 2 (EC 3.1.3.56) (SYNAPTIC INOSITOL-1,4,5-TRISPHOSPHATE 5-
DE PHOSPHATASE 2) (FRAGMENT).
GN SYNJ2 OR KIAA0348.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA Nemoto Y., de Camilli P.;
RT "Characterization of an alternative spliced form of synaptojanin 2.";
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 331-1443 FROM N.A.
RA TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RX Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RX Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC -1- FUNCTION: INOSITOL 5-PHOSPHATASE WHICH MAY BE INVOLVED IN DISTINCT
CC MEMBRANE TRAFFICKING AND SIGNAL TRANSDUCTION PATHWAYS.
CC -1- CATALYTIC ACTIVITY: D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + H(2)O -
CC D-MYO-INOSITOL 1,4-BISPHOSPHATE + PHOSPHATE.
CC -1- SUBUNIT: BINDS TO GRB2.
CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY ASSOCIATED WITH THE
CC PARTICULATE FRACTIONS (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE C-TERMINAL PROLINE-RICH REGION MEDIATES BINDING ONLY
CC TO THE SH3 DOMAIN-CONTAINING PROTEIN GRB2.
CC -1- SIMILARITY: IN THE CENTRAL SECTION, BELONGS TO THE INOSITOL-1,4,5-
CC TRISPHOSPHATE 5-PHOSPHATASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 RAC1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: AF039945; AAD02178.1; -.
DR EMBL: AB002346; BAA20805.2; -.
DR InterPro: IPR000300; -.

```

```

DR InterPro; IPR000504; -
DR InterPro; IPR002013; -
DR Pfam; PF00783; IPIC; 1.
DR PROSITE; PS0102; RRM; 1.
KW Hydrolyase; Alternative splicing; RNA-binding; Multigene family.
FT NO_TERM 1 1
FT DOMAIN 1 396 SACL (BY SIMILARITY).
FT DOMAIN 397 ? CATALYTIC (BY SIMILARITY).
FT DOMAIN 836 915 RNA-BINDING (RRM).
FT DOMAIN 1058 1061 POLY-PRO.
SQ SEQUENCE 1443 AA; 159953 MW; 6C5DAE90FCC1B02B CRC64;

Query Match 52.8%; Score 47.5; DB 1; Length 1443;
Best Local Similarity 71.4%; Pred. No. 1;le+02;
Matches 10; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 3 RRRPPYLP-RRPP 15
DB 1048 RRRPPQPPQRRPP 1061

RESULT 33
BCT5_BOVIN STANDARD; PRT; 176 AA.
AC P19660;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BACTENECCIN 5 PRECURSOR (BAC5) (PR-42).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Neutrophils;
RZ MEDLINE-93107055; PubMed-8416958;
RA Zanetti M., del Sal G., Storici P., Schneider C., Romeo D.;
RT "The CDNA of the neutrophil antiheliotic Bac5 predicts a pro-sequence
RT homologous to a cysteine proteinase inhibitor that is common to other
RT neutrophil antiheliotics."
RL J. Biol. Chem. 268:522-526(1993).
RN [2]
RP SEQUENCE OF 131-173.
RC TISSUE-Neutrophils;
RZ MEDLINE-91035404; PubMed-2229048;
RA Frank R.W., Gennaro R., Schneider K., Przybylski M., Romeo D.;
RT "Amino acid sequences of two proline-rich bactenecins. Antimicrobial
RT peptides of bovine neutrophils."
RL J. Biol. Chem. 265:18871-18874(1990).
RN [3]
RP STRUCTURE BY NMR OF 131-173.
RZ MEDLINE-96183292; PubMed-8605180;
RA Raj P.A., Marcus E., Edgerton M.;
RT "Definition of an active fragment and poly(L-proline) II
RT conformation for candidal activity of bactenecin 5."
RL Biochemistry 35:4314-4325(1996).
RN [4]
RP CHARACTERIZATION.
RZ MEDLINE-96300243; PubMed-8706679;
RA Storici P., Tossi A., Lenarcic B., Romeo D.;
RT "Purification and structural characterization of bovine
RT cathelicidins, precursors of antimicrobial peptides."
RL Eur. J. Biochem. 238:769-776(1996).
CC -1- FUNCTION: EXERTS, IN VITRO, A POTENT ANTIMICROBIAL ACTIVITY.
CC PROBABLY DUE TO AN IMPAIRMENT OF THE FUNCTION OF THE RESPIRATORY
CC CHAIN AND OF ENERGY-DEPENDENT ACTIVITIES IN THE INNER MEMBRANE
CC OF SUSCEPTIBLE MICROORGANISMS.
CC -1- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
CC -1- DOMAIN: BAC5 SEQUENCE CONSISTS ALMOST EXCLUSIVELY OF X-P-P-Y
CC REPEATS.

```

```

CC -1- PM: ELASTASE IS RESPONSIBLE FOR ITS MATURATION.
CC -1- MASS SPECTROMETRY: MW=16570; MW_ERR=1; METHOD-ELECTROSPRAY;
CC RANGE=30-173.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; I02650; AAA30404.1; -
DR PIR; B36589; B36589.
DR PIR; A45328; A45328.
DR InterPro; IPR001894; -
DR Pfam; PF00666; Cathelicidins; 1.
DR PROSITE; PS00946; CATHELICIDINS_1; 1.
DR PROSITE; PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Repeat; Signal; Amidation.
FT SIGNAL 1 29
FT PROPEP 30 130 POTENTIAL.
FT PEPTIDE 131 173 BACTENECCIN 5.
FT PROPEP 174 176 REMOVED IN MATURE FORM.
FT MOD_RES 30 30 PYROLIDONE CARBOXYLIC ACID.
FT DISULFID 85 96 BY SIMILARITY.
FT DISULFID 107 124 BY SIMILARITY.
FT MOD_RES 173 173 AMIDATION (G-174 PROVIDE AMIDE GROUP).
FT CONFLICT 170 171 GP -> R (IN REF. 2).
SQ SEQUENCE 176 AA; 20030 MW; D2DM881C05929415 CRC64;

Query Match 52.2%; Score 47; DB 1; Length 176;
Best Local Similarity 62.5%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 2 RRP--RPPYLRPP 15
DB 137 RRRPPRRPPYPPRRPP 152

RESULT 34
PRP2_MOUSE STANDARD; PRT; 261 AA.
AC P05142;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROLINE-RICH PROTEIN MP-2 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RZ MEDLINE-86059475; PubMed-2999141;
RA Ann D.K., Carlson D.M.;
RT "The structure and organization of a proline-rich protein gene of a
RT mouse multigene family."
RL J. Biol. Chem. 260:15863-15872(1985).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12099; AAA40004.1; -
DR MGD; MGI:97773; Ptp.

```

KW Repeat; Saliva; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 261 PROLINE-RICH PROTEIN MP-2.
 SQ SEQUENCE 261 AA; 26034 MW; 36E13BA7387F47D4 CRC64;

Query Match
 Best Local Similarity 52.2%; Score 47; DB 1; Length 261;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 PPPPYLRPRPP 15
 1 1 1 1 1 1 1 1
 DB 150 PPPPAGPQPRPP 161

RESULT 35
 PRP3_MOUSE STANDARD; PRT; 296 AA.

AC P05143;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE PROLINE-RICH PROTEIN MP-3 (FRAGMENT).
 GN PRP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86059475; PubMed=2999141;
 RA Ann D.K., Carlson D.M.;
 RT "The structure and organization of a proline-rich protein gene of a mouse multigene family.";
 RL J. Biol. Chem. 260:15863-15872(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M12100; AAA4005.1; -
 DR MGD: MGI:97773; PIP.
 KW Repeat; Saliva.
 FT NON_TER 1 1
 SQ SEQUENCE 296 AA; 29521 MW; 7F1A6824B8AF3269 CRC64;

Query Match
 Best Local Similarity 52.2%; Score 47; DB 1; Length 296;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 PPPPYLRPRPP 15
 1 1 1 1 1 1 1 1
 DB 143 PPPPAGPQPRPP 154

RESULT 36
 MEFA_HUMAN STANDARD; PRT; 507 AA.

AC O02078; Q14223; Q14224;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MYOCYTE-SPECIFIC ENHANCER FACTOR 2A (SERUM RESPONSE FACTOR-LIKE
 DE PROTEIN 1).
 GN MEFA2 OR MEF2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart, and Skeletal muscle;
 RX MEDLINE=92387551; PubMed=1516833;
 RA Yu Y.-T., Breitbart R.E., Smoot L.B., Lee Y., Mahdavi V.,
 RA Nadal-Ginard B.;
 RT "Human myocyte-specific enhancer factor 2 comprises a group of
 RT tissue-restricted MADS box transcription factors.";
 RL Genes Dev. 6:1783-1798(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92084105; PubMed=1748287;
 RA Pollock R., Treisman R.;
 RT "Human SRF-related proteins: DNA-binding properties and potential
 RT regulatory targets.";
 RL Genes Dev. 5:2327-2341(1991).
 RN (3)
 RP REVISIONS.
 RA Treisman R.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A. (RSRRC4 AND RSRRC9).
 RA Suzuki E., Lowry J., Sonoda G., Testa J.R., Walsh K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSCRIPTION FACTOR WHICH BINDS SPECIFICALLY TO THE
 CC MEF2 ELEMENT PRESENT IN THE REGULATOR REGIONS OF MANY MUSCLE-
 CC SPECIFIC GENES. ACTIVATES TRANSCRIPTION VIA THIS ELEMENT. MAY BE
 CC INVOLVED IN MUSCLE-SPECIFIC AND/OR GROWTH FACTOR-RELATED
 CC TRANSCRIPTION.
 CC -1- SUBUNIT: HETERODIMER.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; MEF2 (SHOWN HERE), MEFA, RSRRC4
 CC AND RSRRC9; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: MEF2 AND MEFA ISOFORMS ARE EXPRESSED ONLY IN
 CC SKELETAL AND CARDIAC MUSCLE AND IN THE BRAIN WHILE THE RSRRC4 AND
 CC RSRRC9 ISOFORMS ARE EXPRESSED IN ALL TISSUES EXAMINED.
 CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
 CC FACTORS. MEF2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X68505; CAA48517.1; -
 DR EMBL: X63381; CAA44979.1; -
 DR EMBL: X68503; CAA48516.1; -
 DR EMBL: U49020; AAB17195.1; -
 DR EMBL: U44889; AAB17195.1; JOINED.
 DR EMBL: U49012; AAB17195.1; JOINED.
 DR EMBL: U49015; AAB17195.1; JOINED.
 DR EMBL: U49016; AAB17195.1; JOINED.
 DR EMBL: U49017; AAB17195.1; JOINED.
 DR EMBL: U49018; AAB17195.1; JOINED.
 DR EMBL: U49019; AAB17195.1; JOINED.
 DR EMBL: U49020; AAB17195.1; -
 DR EMBL: U44889; AAB17196.1; JOINED.
 DR EMBL: U49012; AAB17196.1; JOINED.
 DR EMBL: U49013; AAB17196.1; JOINED.
 DR EMBL: U49015; AAB17196.1; JOINED.
 DR EMBL: U49016; AAB17196.1; JOINED.
 DR EMBL: U49017; AAB17196.1; JOINED.
 DR EMBL: U49018; AAB17196.1; JOINED.
 DR EMBL: U49019; AAB17196.1; JOINED.
 DR HSSP: P11831; ISSR.
 DR MIM: 600660; -
 DR InterPro: IPR002100; -
 DR Pfam: PF00319; SRF-TF; 1.

DR PRINTS: PR00404; MADSDOMAIN.
 DR PROSITE; PS00350; MADS_BOX.1; 1.
 DR PROSITE; PS00666; MADS_BOX.2; 1.
 KM Transcription regulation; Nuclear protein; DNA-binding; Activator;
 KM Multigene family; Alternative splicing; Phosphorylation.
 FT DOMAIN 3 57
 FT DNA_BIND 58 86 MADS.
 FT DOMAIN 4 31 MEF2-TYPE (POTENTIAL).
 FT DOMAIN 141 186 LYS-RICH (BASIC).
 FT DOMAIN 420 446 SER/THR-RICH.
 FT MOD_RES 289 289 GLN/PRO-RICH.
 FT VARSPPLIC 87 132 PHOSPHORYLATION (BY CK2) (POTENTIAL).
 FT ALNKKHRCDSPPDTSVLPHTHEKYNKNEEDNNMR
 FT NHKIA -> TURKGLNGESPDADYFPHSPLESEDFRSL
 FT NEDSDFIKRGP (IN ISOFORM MEFA AND
 FT RSREC9).
 FT VARSPPLIC 288 295 MISSING (IN ISOFORM RSREC4 AND ISOFORM
 FT RSREC9).
 FT VARSPPLIC 420 421 MISSING (IN ISOFORM RSREC4 AND ISOFORM
 FT RSREC9).
 FT CONFLICT 430 430 MISSING (IN REF. 4).
 FT SEQUENCE 507 AA; 54811 MW; 362BA4FBCC792CE2 CRC64;

Query Match 52.2%; Score 47; DB 1; Length 507;
 Best Local Similarity 46.7%; Pred. No. 48;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPRPYLPFRPP 15
 DB 428 QQQPPPPPPPPPP 442

RESULT 37
 APG_ARATH ID APG_ARATH STANDARD; PRT: 534 AA.
 AC P04602;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE ANTER-SPECIFIC PROLINE-RICH PROTEIN APG. PRECURSOR.
 GN APG.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94004980; PubMed-8401599;
 RA Roberts M.R., Foster G.D., Blundell R.P., Robinson S.W., Kumar A.,
 RA Draper J., Scott R.;
 RT Arabidopsis and sporophytic expression of an anther-specific
 RT Arabidopsis thaliana gene.";
 RL Plant J. 3:111-120(1993).
 CC -1- TISSUE SPECIFICITY: FOUND IN SPOROPHYTIC AND GAMETOPHYTIC CELL
 CC TYPES IN THE ANTHER, ONLY IN MALE FERTILE PLANTS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN MALE GAMETOGENESIS, DURING
 CC MICROSPORE DEVELOPMENT. HIGHER EXPRESSION IS FOUND DURING
 CC MICROSPORE MITOSIS WITH A DRAMATIC DECLINE DURING POLLEN
 CC MATURATION.
 CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X60377; CAA42925.1; -
 CC PIR; S21961; S21961.
 DR InterPro; IPR001087; -

DR Pfam: PF00657; Lipase_GDSL. 1.
 DR PROSITE; PS01098; LIPASE_GDSL_SER. 1.
 KW Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 534
 FT ACT_SITE 211 211 ANTER-SPECIFIC PROLINE-RICH PROTEIN APG.
 FT ACT_SITE 511 511 BY SIMILARITY.
 FT SEQUENCE 534 AA; 57967 MW; 744CAD3B08CC482E CRC64;

Query Match 52.2%; Score 47; DB 1; Length 534;
 Best Local Similarity 53.8%; Pred. No. 51;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 RPRPYLPFRPP 15
 DB 135 KPAPPAPKPPAP 147

RESULT 38
 IRS1_HCMVA ID IRS1_HCMVA STANDARD; PRT: 846 AA.
 AC P09715;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN HORFL.
 GN IRS1.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OC NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87169717; PubMed=3031311;
 RX Weston K., Barrell B.G.;
 RT "Sequence of the short unique region, short repeats, and part of the
 RT long repeats of human cytomegalovirus.";
 RL J. Mol. Biol. 192:177-208(1986).
 RN [2]
 RP COMPLETE GENOME.
 RX MEDLINE-90269039; PubMed=2161319;
 RX Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchinson C.A. III, Kourzides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X17403; CAA35311.1; -
 CC PIR; X04650; CAA28312.1; -
 DR PIR; C26078; QQBEC3.
 DR HSSP; PA1249; S09914.
 KW Hypothetical protein.
 FT CARBOHYD 75 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 846 AA; 91048 MW; 622B6CFE211674BD CRC64;

Query Match 52.2%; Score 47; DB 1; Length 846;
 Best Local Similarity 72.7%; Pred. No. 79;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;


```

QY 3 RRRPPYLRPR 13
   ||||| |||:
Db 824 RRRPPRPRR 834

RESULT 39
BO12 YEAST STANDARD; PRT: 1040 AA.
ID BO12 YEAST
AC P399699
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE BO12 PROTEIN (BEBI PROTEIN).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN RP SEQUENCE FROM N.A.
RA Matsui Y., Matsui R., Toh E. A.;
RL Submitted (XX-1994) to the EMBL/GenBank/DBJ databases.
[2]
RC STRAIN=8288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Bero A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE BEBI PROTEIN. INVOLVED IN BUD FORMATION.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38310; BAA07427.1; -
DR EMBL: U18916; AAC03212.1; -
DR SGD: S0000916; BO12.
DR InterPro: IPR001452; -
DR InterPro: IPR001660; -
DR InterPro: IPR001849; -
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00536; SAM; 1.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50003; PH DOMAIN; 1.
DR Cytoskeleton; SH3 domain.
KW DOMAIN 43 107 SH3.
FT DOMAIN 768 887 PH.
FT CONFLICT 733 733 G->A (IN REF. 1).
SQ SEQUENCE 1040 AA; 11568 MW; 2DC635B1A34E7479 CRC64;

```

Query Match 52.2%; Score 47; DB 1; Length 1040;
 Best local Similarity 53.3%; Pred. No. 95;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRPPYLRPR 15
 ||||| |||:
 Db 440 RRAKPPSPSPAP 454

```

RESULT 40
ID CASR_BOVIN STANDARD; PRT: 1085 AA.
AC P35384;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE EXTRACELLULAR CALCIUM-SENSING RECEPTOR PRECURSOR (CASR) (PARATHYROID
DE CELL CALCIUM-SENSING RECEPTOR).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN RP SEQUENCE FROM N.A.
RA Tissue=Parathyroid;
RX MEDLINE=94077182; PubMed=8255296;
RA Brown E.M., Gamba G., Riccardi D., Lombardi M., Butters R., Kifor O.,
RA Sun A., Hediger M.A., Lytton J., Hebert S.C.;
RT Cloning and characterization of an extracellular Ca(2+)-sensing
RT receptor from bovine parathyroid.";
RL Nature 366:575-580(1993).
CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S67307; AAB29171.1; -
DR EMBL: S40476; S40476.
DR GCRDB: GCR 0900; -
DR InterPro: IPR000068; -
DR InterPro: IPR000337; -
DR InterPro: IPR001828; -
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF receptor; 1.
DR PRINTS: PR00248; GPCRMR.
DR PRINTS: PR00592; CASSENSING.
DR PROSITE: PS00979; G-PROTEIN_RECEP_F3_1; 1.
DR PROSITE: PS00980; G-PROTEIN_RECEP_F3_2; 1.
DR PROSITE: PS00981; G-PROTEIN_RECEP_F3_3; 1.
DR PROSITE: PS0259; G-PROTEIN_RECEP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 1085
FT DOMAIN 20 613
FT TRANSMEM 614 636
FT TRANSMEM 637 650
FT TRANSMEM 651 671
FT TRANSMEM 672 682
FT TRANSMEM 683 701
FT TRANSMEM 702 725
FT TRANSMEM 726 746
FT TRANSMEM 747 770
FT TRANSMEM 771 793
FT TRANSMEM 794 806
FT TRANSMEM 807 829
FT TRANSMEM 830 837
FT TRANSMEM 838 863
FT DOMAIN 864 1085
FT CARBOHYD 91 91

```

EXTRACELLULAR CALCIUM-SENSING RECEPTOR.
 I (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 II (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 III (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 IV (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 V (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VI (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VII (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 1085 AA; 121170 MW; 5D6DDE8C9CD13E47 CRC64;

Query Match
Best Local Similarity 52.2%; Score 47; DB 1; Length 1085;
Matches 7; Conservative 58.3%; Pred. No. 99; Mismatches 3; Indels 0; Gaps 0;

OY 3 RPPRYLP RPP 14
DB 950 QRPSTPPOP 961

RESULT 41
EXTN_MAIZE STANDARD; PRT; 267 AA.
ID P14918;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
GN HRGP.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. W64A, AND CV. E41;
RA Stiefel V., Perez-Gran L., Albericio F., Giralt E., Ruiz-Avila L.,
RA Ludevid M.D., Puligomenech P.;
RT "Molecular cloning of cDNAs encoding a putative cell wall protein from
RT Zea mays and immunological identification of related polypeptides.";
RL Plant Mol. Biol. 11:483-493(1988).
CC -1- FUNCTION: STRUCTURAL COMPONENT IN PRIMARY CELL WALL.
CC -1- TISSUE SPECIFICITY: MAINLY IN THE COLEOPTILE NODE AND ROOT TIP.
CC -1- PTM: EXTENSIN CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
CC GLYCOSYLATED.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X13499; CAA31854.1; -
DR EMBL: X13506; CAA31860.1; -
DR EMBL: M36912; AAA33455.1; -
DR EMBL: M36913; AAA33456.1; -
DR EMBL: M36914; AAA33457.1; -
DR PIR: S08314; S08314.
DR Matedb: 17152; -
DR InterPro: IPR002965; -
DR PRINTS: PRO1217; PRICEXTENSIN.
KW Repeat: Cell wall. Glycoprotein. Signal. Structural protein;
KW Hydroxylation.
FT SIGNAL 1 ?
FT CHAIN 267 ?
FT DOMAIN 18 253 EXTENSIN.
FT REPEAT 18 33 HIGHLY REPETITIVE ZONE.
FT REPEAT 34 54

```

```

FT REPEAT 55 70
FT REPEAT 71 91
FT REPEAT 92 107
FT REPEAT 108 128
FT REPEAT 129 144
FT REPEAT 145 160
FT REPEAT 161 179
FT REPEAT 180 195
FT REPEAT 196 211
FT REPEAT 212 232
FT REPEAT 233 253
FT DOMAIN 261 265
FT VARIANT 245 245
FT VARIANT 261 261
SO SEQUENCE 267 AA; 28349 MW; A6F406F4645FEBCB CRC64;

Query Match
Best Local Similarity 51.7%; Score 46.5; DB 1; Length 267;
Matches 8; Conservative 57.1%; Pred. No. 30; Mismatches 3; Indels 1; Gaps 1;

OY 3 RPPRYLP RPP 15
DB 31 KRPPTYTPSPKP 44

RESULT 42
EXTN_DICDI STANDARD; PRT; 424 AA.
ID SP60_DICDI
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE SPORE COAT PROTEIN SP60 PRECURSOR.
GN COFC.
OC Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90097939; PubMed=2601718;
RA Eosnaukh K.L., Loomis W.F.; SP70 of Dictyostelium discoideum.";
RT "Spore coat genes SP60 and SP70 of Dictyostelium discoideum.";
RL Mol. Cell. Biol. 9:5215-5218(1989).
RN [2]
RP SEQUENCE OF 1-185 FROM N.A.
RX MEDLINE=90299131; PubMed=2163344;
RA Haberstroth L., Firtel R.A.;
RT "A spatial gradient of expression of a CAMP-regulated prespore
RT cell-type-specific gene in Dictyostelium.";
RL Genes Dev. 4:596-612(1990).
RN [3]
RP REVISIONS.
RX MEDLINE=91014695; PubMed=2170814;
RA Widdowson D.C.C., Profitilt J.A., Jagger P.S., Richards A.J.,
RA Hames B.D.;
RT "Developmental expression and characterization of the gene encoding
RT spore coat protein SP60 in Dictyostelium discoideum.";
RL Mol. Microbiol. 4:951-960(1990)
CC -1- SIMILARITY: CONTAINS 3 PRESPORE MOTIFS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51892; CAA36174.1; -
DR EMBL: M26239; AAA33251.1; -
DR EMBL: X52105; CAA36325.1; -
DR PIR: A33485; A33485.

```

Query	Match	Best local	Similarity	Score	DB	Length	Matches	Mismatches	Indels	Gaps
QY	3	RRPPRYLP	RRPPRP	15	51.7%;	46.5;	1;	2;	1;	1
Db	172	RPR-PHRRP	PHRRP	183	69.2%;	47;	1;	2;	1;	1
<p>RESULT 43</p> <p>POLY_SINDV STANDARD; PRT; 1245 AA.</p> <p>AC P03316; Q88870; Q88871; Q88872; Q88873; Q88874;</p> <p>DT 21-JUL-1986 (Rel. 01, Created)</p> <p>DT 21-JUL-1986 (Rel. 01, Last sequence update)</p> <p>DT 15-DEC-1998 (Rel. 37, Last annotation update)</p> <p>DE STRUCTURAL POLYPROTEIN (P130) [CONTAINS: COAT PROTEIN C (EC 3.4.21.-); SPIKE GLYCOPROTEIN E3, E2 AND E1; 6 KDA PEPTIDE].</p> <p>OS Sindbis virus (strains HRSP and HRLP).</p> <p>OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.</p> <p>OC NCBI_TaxID-11034;</p> <p>BN [1]</p> <p>BN SEQUENCE FROM N.A.</p> <p>RP STRAIN-HRSP;</p> <p>RC MEDLINE-84148439; PubMed-6322438;</p> <p>RA Strauss E.G., Rice C.M., Strauss J.H.,</p> <p>RT "Complete nucleotide sequence of the genomic RNA of Sindbis virus,"</p> <p>RL Virology 133:92-110(1984).</p> <p>RN [2]</p> <p>RP SEQUENCE FROM N.A.</p> <p>RC STRAIN-HRSP;</p> <p>RC MEDLINE-81223817; PubMed-6941270;</p> <p>RA Rice C.M., Strauss J.H.;</p> <p>RT "Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced</p> <p>RL sequence of the encoded virus structural proteins,"</p> <p>RL Proc. Natl. Acad. Sci. U.S.A. 78:2062-2066(1981).</p> <p>RN [3]</p> <p>RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 107-264.</p> <p>RC MEDLINE-96428684; PubMed-8831786;</p> <p>RA Choi H.-K., Lee S., Zhang Y.-P., McKinney B.R., Wengler G.,</p> <p>RT Rossmann M.G., Kuhn R.J.;</p> <p>RT "Structural analysis of Sindbis virus capsid mutants involving</p> <p>RT assembly and catalysis,"</p> <p>RL J. Mol. Biol. 262:151-167(1996).</p> <p>RN [4]</p> <p>RP ERRATUM.</p> <p>RA Choi H.-K., Lee S., Zhang Y.-P., McKinney B.R., Wengler G.,</p> <p>RA Rossmann M.G., Kuhn R.J.;</p> <p>RL J. Mol. Biol. 266:633-634(1997).</p> <p>RL -1- PUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.</p> <p>CC -1- PWM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.</p> <p>CC -1- MISCELLANEOUS: THE 6 KDA POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL</p> <p>SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL</p>										

```

CC      HEMAGGLUTININ.
CC      -1- MISCELLANEOUS: THE STRAIN HRPD SEQUENCE IS SHOWN.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to licensese@isb-sib.ch).
CC      -----
CC      EMBL: V01403; CAA24684.1; -.
CC      DR EMBL: J02363; AAA96976.1; -.
CC      DR PIR: A03916; VHWVB.
CC      DR PIR: B03916; VHWV2.
CC      DR PDB: 1KXA; 08-NOV-96.
CC      DR PDB: 1KXB; 08-NOV-96.
CC      DR PDB: 1KXC; 08-NOV-96.
CC      DR PDB: 1KXD; 08-NOV-96.
CC      DR PDB: 1KXE; 08-NOV-96.
CC      DR PDB: 1KXF; 07-DEC-96.
CC      DR PDB: 2SNW; 08-APR-98.
CC      DR PDB: 1WYK; 28-APR-98.
CC      DR MEROPS: S03.001; -.
CC      DR InterPro: IPR000930; -.
CC      DR InterPro: IPR000936; -.
CC      DR InterPro: IPR001836; -.
CC      DR InterPro: IPR002533; -.
CC      DR InterPro: IPR002548; -.
CC      DR Pfam: PF01589; Alpha_E1_glycop; 1.
CC      DR Pfam: PF00943; Alpha_E2_glycop; 1.
CC      DR Pfam: PF01563; Alpha_E3_glycop; 1.
CC      DR Pfam: PF00944; Alpha_core; 1.
CC      DR PRINTS: PR00798; TCGAVIRIN.
CC      KW Coat protein; Polypeptide; Transmembrane; Glycoprotein; Hydrolase;
KW      Serine protease; 3d-structure.
KW      FT CHAIN 1 264 COAT PROTEIN C (CAPSID PROTEIN C).
KW      FT CHAIN 265 328 SPIKE GLYCOPROTEIN E3.
KW      FT CHAIN 329 751 SPIKE GLYCOPROTEIN E2.
KW      FT CHAIN 752 806 6 KDA PEPTIDE.
KW      FT CHAIN 807 1245 SPIKE GLYCOPROTEIN E1.
KW      FT ACT_SITE 141 141 CHARGE RELAY SYSTEM (BY SIMILARITY).
KW      FT ACT_SITE 147 147 CHARGE RELAY SYSTEM (BY SIMILARITY).
KW      FT ACT_SITE 215 215 CHARGE RELAY SYSTEM (BY SIMILARITY).
KW      FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
KW      FT CARBOHYD 524 524 N-LINKED (GLCNAC. . .) (POTENTIAL).
KW      FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
KW      FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
KW      FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
KW      FT VARIANT 333 333 D -> G (IN STRAIN HRPD).
KW      FT VARIANT 351 351 V -> E (IN STRAIN HRPD).
KW      FT VARIANT 919 919 D -> V (IN STRAIN HRPD).
KW      SQ SEQUENCE 1245 AA; 136765 MW; B77C18131703F1E6 CRC64;

Query Match 51.7%; Score 46.5; DB 1; Length 1245;
Best Local Similarity 69.2%; Pred. No. 1.3e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 3; Gaps 1.

QY 3 RRPPLYLRPRPP 15
   ||| |||||
Db 59 RPPQ---RPPRP 68

RESULT 44
Y028_NPVOP STANDARD; PRT; 138 AA.
AC 010295;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 16.5 KDA PROTEIN (OHF40).

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:05:20 ; Search time 47.47 Seconds
(without alignments)
41.807 Million cell updates/sec

Title: US-09-276-868-3
Perfect score: 90
Sequence: 1 RRRPPPLPRRPP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	61	67.8	336	14	068405	068405 human cytom
2	58	64.4	164	6	P79361	P79361 ovls aries
3	58	64.4	190	6	Q9X5Q9	Q9X5Q9 capra hircu
4	58	63.4	224	6	019031	019031 ovls aries
5	57	63.3	200	2	Q9RK54	Q9RK54 streptomyce
6	57	63.3	361	2	Q9XCG4	Q9XCG4 mycobacteri
7	56	62.2	212	2	008306	008306 nocardioid
8	55	61.1	212	14	041980	041980 murid herpe
9	55	61.1	437	5	017015	017015 caenorhabd
10	54.5	60.6	301	10	041848	041848 zea mays (m
11	54.5	60.6	2635	14	P88955	P88955 kaposi's sa
12	54.5	60.6	2635	14	040942	040942 kaposi's sa
13	54	60.0	359	5	Q9XZT0	Q9XZT0 drosophila
14	54	60.0	427	5	044582	044582 caenorhabd
15	54	60.0	955	4	Q9Y2W1	Q9Y2W1 homo sapien
16	53	58.9	168	10	Q9SM77	Q9SM77 oryza sativ
17	53	58.9	184	5	023291	023291 caenorhabd
18	53	58.9	333	10	Q9XIZ3	Q9XIZ3 oryza sativ
19	53	58.9	381	6	Q9GKN8	Q9GKN8 bos taurus

20	53	58.9	428	10	023370	023370 arabidopsis
21	53	58.9	491	10	082066	082066 solanum tub
22	53	58.9	520	10	Q9LV14	Q9LV14 arabidopsis
23	52.5	58.3	602	14	Q66852	Q66852 fowl adenov
24	52	57.8	439	10	Q42421	Q42421 beta vulgar
25	52	57.8	467	10	Q39184	Q39184 arabidopsis
26	52	57.8	1006	10	Q9LM01	Q9LM01 arabidopsis
27	52	57.8	1091	5	Q9W126	Q9W126 drosophila
28	51	56.7	94	5	Q917F1	Q917F1 drosophila
29	51	56.7	409	5	Q9U0Z7	Q9U0Z7 leishmania
30	51	56.7	417	5	Q9V4V1	Q9V4V1 drosophila
31	51	56.7	1307	5	Q9V4J6	Q9V4J6 drosophila
32	51	56.7	3325	14	Q9IBT9	Q9IBT9 turkey herp
33	50.5	56.1	151	10	Q9FI95	Q9FI95 arabidopsis
34	50.5	56.1	249	10	Q24102	Q24102 medicago tr
35	50.5	56.1	282	10	Q9FP63	Q9FP63 oryza sativ
36	50.5	56.1	1479	11	088399	088399 mus musculu
37	50	55.6	96	14	Q9Q7W0	Q9Q7W0 tt virus. o
38	50	55.6	192	2	P74661	P74661 synechocyst
39	50	55.6	344	14	069260	069260 bovine herp
40	50	55.6	383	10	Q9SVY1	Q9SVY1 arabidopsis
41	50	55.6	409	10	Q9SBM1	Q9SBM1 volvox cart
42	50	55.6	421	11	060491	060491 cavia porce
43	50	55.6	464	10	Q9FF14	Q9FF14 arabidopsis
44	50	55.6	593	10	Q9MA04	Q9MA04 arabidopsis
45	50	55.6	596	5	Q9VNA6	Q9VNA6 drosophila

ALIGNMENTS

RESULT 1
068405 ID Q68405 PRELIMINARY; PRT; 336 AA.
AC 068405;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ORF UL151.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOLEDO;
RX MEDLINE=9609416; PubMed=8523595;
RA Cha T.A., Tom E., Kemble G.W., Duke G.M., Mocarski E.S., Spaete R.R.;
RT "Human cytomegalovirus clinical isolates carry at least 19 genes not
found in laboratory strains."
RL J. Virol. 70:78-83(1996).
DR EMBL: U33331; AAA85892.1; -
SQ SEQUENCE 336 AA; 35116 MW; 9F865E5019F69D0C CRC64;

Query Match 67.8%; Score 61; DB 14; Length 336;
Best Local Similarity 78.6%; Pred. No. 0.76;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRRPPPLPRRPP 15
DB 279 RRRPPPLPRRPP 292
RESULT 2
ID P79361 PRELIMINARY; PRT; 164 AA.
AC P79361;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE 7.5 KDA BACTENEICIN PRECURSOR (FRAGMENT).
GN BAC7.5.

OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 NCBI_TaxID=9940;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RL Hutter K.M., Mahoney M.M.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: EXERTS, IN VITRO, A POTENT ANTIMICROBIAL ACTIVITY.
 CC PROBABLY DUE TO AN IMPAIRMENT OF THE FUNCTION OF THE RESPIRATORY
 CC CHAIN AND OF ENERGY-DEPENDENT ACTIVITIES IN THE INNER MEMBRANE OF
 CC SUSCEPTIBLE MICROORGANISMS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 DR EMBL: U60598; AAB49713.1; -;
 DR InterPro: IPR001894; -;
 DR Pfam: PF00666; Cathelicidins; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 FT NON_TER 164 164
 FT SEQUENCE 164 AA; 18642 MW; E3BFC871F6AEB9A CRC64;

Query Match 64.4%; Score 58; DB 6; Length 164;
 Best Local Similarity 78.6%; Pred. No. 0.99;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPRPYLRPRP 14
 | | | | | | | | | |
 Db 132 RLRRPRRLPRPR 145

RESULT 3
 O9XS09 PRELIMINARY; PRT; 190 AA.
 AC O9XS09;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE BAC7.5 PROTEIN.
 GN BAC7.5.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 NCBI_TaxID=9925;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RA Zhao C., Nguyen T., Brogden K., Lehrer R.;
 RA "cDNA cloning of goat cathelin related peptides."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ24125; CAB45523.1; -;
 DR InterPro: IPR001894; -;
 DR Pfam: PF00666; Cathelicidins; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 FT CHAIN 131 190
 FT SEQUENCE 190 AA; 21835 MW; D13305EF16875F4F CRC64;

Query Match 64.4%; Score 58; DB 6; Length 190;
 Best Local Similarity 78.6%; Pred. No. 1.1;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPRPYLRPRP 14
 | | | | | | | | | |
 Db 132 RLRRPRRLPRPR 145

RESULT 4
 O19031

ID O19031 PRELIMINARY; PRT; 224 AA.
 AC O19031;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE BACTINECIN 11 PRECURSOR.
 GN BAC11.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 NCBI_TaxID=9940;
 [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=98121317; PubMed=9461419;
 RA Hutter K.M., Lambeth M.R., Burkin H.R., Broad T.E.;
 RT "Localization and genomic organization of sheep antimicrobial peptides
 RT genes.";
 RL Gene 206:85-91(1998).
 CC -1- FUNCTION: ANTIMICROBIAL PEPTIDE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 DR EMBL: U77049; AAB62000.1; -;
 DR EMBL: U77046; AAB62000.1; JOINED.
 DR EMBL: U77047; AAB62000.1; JOINED.
 DR EMBL: U77048; AAB62000.1; JOINED.
 DR InterPro: IPR001894; -;
 DR Pfam: PF00666; Cathelicidins; 1.
 DR ProDom: PD001838; -; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 KW Signal; Antibiotic.
 FT SIGNAL 1 29
 FT PROPEP 30 130
 FT CHAIN 131 224
 FT MOD_RES 30 30
 FT DISULFID 85 96
 FT DISULFID 107 124
 FT SEQUENCE 224 AA; 25669 MW; 6AEAB1256AC76FC CRC64;

Query Match 64.4%; Score 58; DB 6; Length 224;
 Best Local Similarity 78.6%; Pred. No. 1.3;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPRPYLRPRP 14
 | | | | | | | | | |
 Db 132 RLRRPRRLPRPR 145

RESULT 5
 O9RK54 PRELIMINARY; PRT; 200 AA.
 AC O9RK54;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 22.1 KDA PROTEIN.
 GN SCF12.02C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 NCBI_TaxID=1902;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Murphy L., Harris D.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.


```
RM [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA MEDLINE-97000351; PubMed-8843436;
RA Redenbach M., Kleser H.M., Denapelle D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.,
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL117669; CAB56128.1;
KW Hypothetical protein.
SQ SEQUENCE 200 AA; 22076 MW; 0DCBBEC5585803B5 CRC64;

Query Match
Best Local Similarity 76.3%; Score 57; DB 2; Length 200;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPPYLP RPR 13
DB 118 RRRPPPALPRPR 130

RESULT 6
QYXCG4 PRELIMINARY; PRT; 361 AA.
AC 09XCG4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 40.2 KDA PROTEIN.
OS Mycobacterium avium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID-1764;
RN [1]
RP SEQUENCE FROM N.A.
RC Eckstein T.M., Lambert M.L., Brennan P.J., Belisle J.T., Inamine J.M.;
RT "Identification of a gene cluster involved in glycopeptidolipid
RT biosynthesis and of a gene cluster encoding daunorubicin resistance in
RT two strains of Mycobacterium avium serovar 2."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF143772; AAD44199.1;
KW Hypothetical protein.
SQ SEQUENCE 361 AA; 40208 MW; AD01DBE825C1C9EA CRC64;

Query Match
Best Local Similarity 63.3%; Score 57; DB 2; Length 361;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRRPPPYLP RPR 14
DB 32 RRRPPRPAHP RPR 45

RESULT 7
QYXCG4 PRELIMINARY; PRT; 212 AA.
AC 008306;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S21.
OS Nocardioides simplex (Arthrobacter simplex).
OC Bacteria; Firmicutes; Actinobacteria; Pimeleobacter.
OX NCBI_TaxID-2045;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-95319331; PubMed-7596291;
RX STRAIN-IFO12069;
RA Molnar I., Choi K.P., Yamashita M., Murooka Y.;
```

```
RT "Molecular cloning, expression in Streptomyces lividans, and analysis
RT of a gene cluster from Arthrobacter simplex encoding 3-ketosteroid-
RT delta 1-dehydrogenase, 3-ketosteroid-delta 5-isomerase and a
RT hypothetical regulatory protein."
RL Mol. Microbiol. 15:895-905(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-IFO12069;
RA Dzadek J., Yamashita M., Murooka Y.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE TETR/ACR FAMILY OF TRANSCRIPTIONAL REGULATORS.
DR EMBL: Z93338; CAB07542.1;
DR InterPro: IPR001647;
DR Pfam: PF00440; tetr; 1.
DR PRINTS: PR00455; HTHTETR.
DR PROSITE: PS01081; HTM_TETR_FAMILY; 1.
KW DNA-binding; Hypothetical protein; Ribosomal protein;
KW Transcription regulation.
SQ SEQUENCE 212 AA; 22740 MW; F9118E18DDF4E0B2 CRC64;

Query Match
Best Local Similarity 62.2%; Score 56; DB 2; Length 212;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRRPPPYLP RPR 15
DB 83 RRRPPSGQ RPR 97

RESULT 8
QYXCG4 PRELIMINARY; PRT; 212 AA.
AC 041980;
ID 041980;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 21.9 KDA PROTEIN.
GN GAMMAHV.M3.
OS murid herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID-33708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WUMS;
RX MEDLINE-97366649; PubMed-9223479;
RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
RA Dal Canto A.J., Speck S.H.;
RT "Complete sequence and genomic analysis of murine gammaherpesvirus
RT 68."
RL J. Virol. 71:5894-5904(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-WUMS;
RA Latreille P., Wamsley P., Waterston R.H.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97553; AAB6426.1;
KW Hypothetical protein.
SQ SEQUENCE 212 AA; 21911 MW; E066860064282149 CRC64;

Query Match
Best Local Similarity 61.1%; Score 55; DB 14; Length 212;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPPPYLP RPR 15
DB 136 PSPPLPR RPR 147

RESULT 9
QYXCG4 PRELIMINARY; PRT; 212 AA.
```

```

ID 017015 PRELIMINARY; PRT; 437 AA.
AC 017015;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE R3D11.3 PROTEIN.
GN R3D11.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodetidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smailon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Miller N., Beck C., Gibson A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
SQ EMBL: AF022983; AAB69949.1;
SQ SEQUENCE 437 AA; 48915 MW; 39A689FD6745EFD6 CRC64;

```

```

Query Match 61.1%; Score 55; DB 5; Length 437;
Best Local Similarity 75.0%; Pred. NO. 5.8;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 PRPPYLP-PRPP 15
DB 23 PRPPYLP-PRPP 34

```

```

RESULT 10
ID 041848 PRELIMINARY; PRT; 301 AA.
AC 041848;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PROLIN RICH PROTEIN.
GN PRP.
OS Zea mays (maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W64A;
RX MEDLINE=92361259; PubMed=1498600;
RA Jose-Estanyol M., Ruiz-Avila L., Puigdomenech P.;
RT "A maize embryo-specific gene encodes a proline-rich and hydrophobic
RT protein."
RL Plant Cell 4:413-423(1992).

```

```

DR EMBL: X60432; CAA42959.1;
DR HSSP; P24337; IHP.
DR Mendel; 16222; Zeama;1531;16222.
DR InterPro; IPR000528;
DR InterPro; IPR003612;
DR Pfam; PF00279; LTP; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 301 AA; 31647 MW; 884EB70854D28C2E CRC64;

```

```

Query Match 60.6%; Score 54.5; DB 10; Length 301;
Best Local Similarity 71.4%; Pred. NO. 4.8;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
OY 3 PRPPYLP-PRPP 15
DB 149 PRPPYLP-PRPP 162

```

```

RESULT 11
ID P88955 PRELIMINARY; PRT; 2635 AA.
AC P88955;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ORF 64.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhabdovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97121480; PubMed=8962146;
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT (HHV8)."
RL Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97094384; PubMed=8939871;
RA Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway
RT genes by KSHV."
RL Science 274:1739-1744(1996).
DR EMBL: U75698; AAC57149.1;
SQ SEQUENCE 2635 AA; 289688 MW; 00070132EA8139AF CRC64;

```

```

Query Match 60.6%; Score 54.5; DB 14; Length 2635;
Best Local Similarity 68.4%; Pred. NO. 33;
Matches 13; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
OY 2 RRPR--PYLP--PRPP 15
DB 271 RRPRVLPYLPYPTDRPP 289

```

```

RESULT 12
ID 040942 PRELIMINARY; PRT; 2635 AA.
AC 040942;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE ORF 64.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhabdovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.

```

RA MEDLINE=97296220; PubMed=9151804;
 RA Neipel F., Albrecht J.C., Fleckenstein B.;
 RT "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
 human herpesvirus 8: determinants of its pathogenicity?";
 RL J. Virol. 71:4187-4192(1997).
 DR EMBL: U93872; AAB62600.1; -;
 SQ SEQUENCE 2635 AA; 289718 MW; 91DDAD06F7B660A CRC64;

Query Match 60.6%; Score 54.5; DB 14; Length 2635;
 Best Local Similarity 68.4%; Pred. No. 33;
 Matches 13; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 2 RRRP---PPLP---RRPP 15
 ||||| ||| | |||||
 DB 271 RRRPVPVPPYDTRPPRP 289

RESULT 13
 Q9XZT0 PRELIMINARY; PRT; 359 AA.

AC Q9XZT0; Q9W5D1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE EG:34F3.10 PROTEIN (CG13358 PROTEIN).
 GN EG:34F3.10 OR CG13358.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;

NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.

RA STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RN [12]
 RP SEQUENCE FROM N.A.

RC STRAIN-OREGON-R;
 RA Valenti P., Salles C., Campbell L., Glover D.,

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003418; AAF45553.1; ALT_INT.

DR EMBL: AL031583; CAB41346.1; -;
 DR FLYBase; FBgn0026874; EG:34F3.10.

SQ SEQUENCE 359 AA; 38488 MW; 083A6CF2C6F1D74E CRC64;

Query Match 60.0%; Score 54; DB 5; Length 359;
 Best Local Similarity 71.4%; Pred. No. 6.6;
 Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRRPPLPPLP 15
 ||| ||| ||| |||
 DB 167 RRRPPLPPLP 180

RESULT 14
 ID 044582 PRELIMINARY; PRT; 427 AA.

AC 044582;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE F39C12.3 PROTEIN.
 GN F39C12.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 NCBI_TaxID=6239;
 [1]
 SEQUENCE FROM N.A.

RP STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lighting J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Koopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RT Nature 368:32-38(1994).

RL Nature 368:32-38(1994).

RN [12]
 RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
 RA Chisose S., Sansone J.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF039043; AAB94196.1; -;

DR InterPro; IPR000301; -;

DR Pfam; PF00335; transmembrane4; 1.

SQ SEQUENCE 427 AA; 47654 MW; 67288DA03AA6B513 CRC64;

Query Match 60.0%; Score 54; DB 5; Length 427;
 Best Local Similarity 69.2%; Pred. No. 7.6;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RRRPPLPPLP 15
 ||||| :| | | |

Db 338 RRRPPDIPPLPP 350

RESULT 15

Q9Y2W1 ID PRELIMINARY: PRT; 955 AA.

AC 09Y2W1

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)

DE 01-JUN-2000 (TREMBLrel. 14, last annotation update)

DE THYROID HORMONE RECEPTOR-ASSOCIATED PROTEIN COMPLEX COMPONENT TRAP150.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Ito M., Yuan C.-X., Malik S., Gu W., Fondell J.D., Yamamura S., Fu Z.-Y., Zhang Z., Qin J., Koeder R.G.,

RT "Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators."

RL M01. Cell 0:0-0(1999).

DR EMBL; AF117756; AAD22034.1; -

KW Receptor.

SQ SEQUENCE 955 AA; 108693 MW; FF031C3632E1E6A8 CRC64;

Query Match 60.0%; Score 54; DB 4; Length 955; Best Local Similarity 60.0%; Pred. NO. 16; Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 RRRRPPYLPPLPP 15

Db 262 RRRPSPVPKPSPP 276

Q9SM77 ID PRELIMINARY: PRT; 168 AA.

AC 09SM77

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, last annotation update)

DE ZHB0008.1.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. INDICA;

RA Hong G., Zhou B.;

RT "Oryza sativa genomic DNA, chromosome 4, clone: t17804."

DR EMBL; A111265; CAB55416.1; -

SQ SEQUENCE 168 AA; 18668 MW; B9A19B483840F1A3 CRC64;

Query Match 58.9%; Score 53; DB 10; Length 168; Best Local Similarity 66.7%; Pred. NO. 4.5; Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 RRRRPPYLPPLPP 15

Db 56 RRRPDPPLRRRCQP 70

RESULT 17

Q23291 ID PRELIMINARY: PRT; 184 AA.

AC 023291

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)

DE COSMID ZC404.

CN ZC404.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R., Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans."

RT Nature 368:32-38(1994).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Bentley D., Le T.T.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U55363; AAA97967.1; -

SQ SEQUENCE 184 AA; 21649 MW; 6D64D09EA91308F9 CRC64;

Query Match 58.9%; Score 53; DB 5; Length 184; Best Local Similarity 90.0%; Pred. NO. 4.9; Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPYLPPLPP 12

Db 26 RRRPYPYLPPLPP 35

RESULT 18

Q9X1Z3 ID PRELIMINARY: PRT; 333 AA.

AC 09X1Z3

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)

DE 01-MAR-2001 (TREMBLrel. 16, last annotation update)

DE SIMILAR TO ZEA MAYS PRP GENE. (X60432).

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. NIPPONBARE;

RA Sasaki T., Nagamura Y., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC clone: P0680A03."

RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB023482; BAA78762.1; -

DR HSSP; P24337; 1HTP.

DR InterPro; IPR000528; -

DR InterPro; IPR002965; -

DR InterPro; IPR002966; -

DR InterPro: IPR003612: -
 DR Pfam: PF00279; LTP; 1.
 DR PRINTS: PRO1217; PRICHTEXTENS.
 DR PRINTS: PRO1218; PSTEXTENSIN.
 DR SMART: SM00499; AAI; 1.
 SQ SEQUENCE 333 AA; 34823 MW; 268DEC74E20E8194 CRC64;

Query Match 58.9%; Score 53; DB 10; Length 333;
 Best Local Similarity 66.7%; Pred. NO. 8.3;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 RRPPLPRPRP 15
 DB 206 PTPPYPPPTPP 217

RESULT 19
 O9GKN8 PRELIMINARY; PRT: 381 AA.

AC O9GKN8; 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PROLINE ARGinine-RICH END LEUCINE-RICH REPEAT PROTEIN.
 GN PRELP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RN NCBL_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RC TISSUE=ARTICULAR CARTILAGE;
 RX MEDLINE=20576219; PubMed=11007795;
 RA Bengtsson E., Asperger A., Heinegard D., Sommarin Y., Spillmann D.;
 RT "The amino-terminal part of PRELP binds to heparin and heparan
 sulfate.";
 RL J. Biol. Chem. 275:40695-40702(2000).
 CR EMBL: AF163568; AAG23723.1; -;
 SQ SEQUENCE 381 AA; 43682 MW; 23DA99C01B772A0 CRC64;

Query Match 58.9%; Score 53; DB 6; Length 381;
 Best Local Similarity 76.9%; Pred. NO. 9.3;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRPPLPRPRP 14
 DB 25 RRPPLPRPRPRP 37

RESULT 20
 O23370 PRELIMINARY; PRT: 428 AA.

AC O23370; 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE COL. WALL PROTEIN HOMOLOG (CELL WALL PROTEIN LIKE).
 GN AT4G15160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 RN NCBL_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Mambutt R., Pohl T., Terryn N.,
 RA Kreis M., Kavanagh T., Enliar K.D., Kieger M., James R.,
 RA Pulgomech P., Hatzopoulos P., Obermayer B., Duesterhoff A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Schueller C., Chawatzis N.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 297338; CAB10295.1; -;
 DR EMBL: AL161540; CAB78558.1; -;
 DR HSP; P24337; IHP.
 DR Mendel; 26707; AtAtch1531; 26707.

DR InterPro: IPR00528; -;
 DR InterPro: IPR01353; -;
 DR InterPro: IPR03612; -;
 DR Pfam: PF00227; proteasome; 2.
 DR Pfam: PF00279; LTP; 1.
 DR SMART: SM00499; AAI; 1.
 SQ SEQUENCE 428 AA; 45864 MW; D0EACBD34C6717FC CRC64;

Query Match 58.9%; Score 53; DB 10; Length 428;
 Best Local Similarity 61.5%; Pred. NO. 10;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRPPLPRPRP 15
 DB 67 KPPPYICPPPP 79

RESULT 21
 ID02066 PRELIMINARY; PRT: 491 AA.

AC O82066; 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE PROLINE-RICH PROTEIN.
 GN GPP1.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; I;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 RN NCBL_TaxID=4113;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. DESIREE;
 RA Meike U., Renault N., Mueller-Roeber B.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ000997; CAA04449.1; -;
 SQ SEQUENCE 491 AA; 54112 MW; A4E648680845F0C CRC64;

Query Match 58.9%; Score 53; DB 10; Length 491;
 Best Local Similarity 50.0%; Pred. NO. 12;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRPPLPRPRP 15
 DB 306 KPCPPSVKPKPP 319

RESULT 22
 O9LV14 PRELIMINARY; PRT: 520 AA.

AC O9LV14; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 5, p1 CLONE:MRG21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 RN NCBL_TaxID=3702;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-20181125; PubMed-10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RT Tabata S.;
 "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 features of the regions of 3,076,755 bp covered by sixty pl and 7AC
 clones.";
 RL DNA Res. 7:31-63(2000).
 DR EMBL; AB020751; BAA97211.1; -
 DR InterPro: IPR002965; -
 DR PRINTS; PR01217; PRICHTEXTNSN.
 SQ SEQUENCE 520 AA; 53996 MW; BB6652EE0BE93E2D CRC64;

Query Match 58.9%; Score 53; DB 10; Length 520;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 4 PRPPYLRPPRP 15
 378 PRPPYLRPPRP 389

RESULT 23
 Q66852 PRELIMINARY; PRT; 602 AA.
 AC Q66852;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PRE-TERMINAL PROTEIN.
 OS fowl adenovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadonovirus.
 OX NCBI_TaxID=31540;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SEROTYPE 10;
 RA MCCoy R.J., Sheppard M.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Sheppard M., Werner W., McCoy R., Johnson M.A.;
 RL Arch. Virol. 0:0-0(1997).
 DR EMBL; U26220; AAB02174.1; -
 DR EMBL; AF007577; AAB88665.1; -
 DR InterPro: IPR003391; -
 DR Pfam; PF02459; Adeno_terminal; 1.
 SQ SEQUENCE 602 AA; 70392 MW; EB48277B4390894F CRC64;

Query Match 58.3%; Score 52.5; DB 14; Length 602;
 Best Local Similarity 54.2%; Pred. No. 16;
 Matches 13; Conservative 0; Mismatches 2; Indels 9; Gaps 2;

DB 1 RRRPRP-----YLP--RRRP 15
 338 RRRPRPSPQPPPEYLPLEPRFP 361

RESULT 24
 Q42421 PRELIMINARY; PRT; 439 AA.
 AC Q42421;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE CHITINASE PRECURSOR.
 GN CHAI OR CH1.
 OS Beta vulgaris (Sugar beet).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
 CC Caryophyllales; Chenopodiaceae; Beta.

OX NCBI_TaxID=3555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LEAF;
 RX MEDLINE-95170004; PubMed-7865792;
 RA Berglund L., Brundstedt J., Nielsen K.K., Chen Z., Mikkelsen J.D.,
 RT Marcker K.A.;
 "A proline-rich chitinase from Beta vulgaris.";
 RL Plant Mol. Biol. 27:211-216(1995).
 DR EMBL; X81056; CAA56946.1; -
 DR EMBL; X81057; CAA56946.1; JOINED.
 DR EMBL; X79301; CAA55883.1; -
 DR EMBL; A23786; CAA01693.1; -
 DR HSSP; P23951; 2BAA.
 DR Mendel, 1557; Betvu; Chai1; 1557.
 DR InterPro: IPR000726; -
 DR Pfam; PF00182; Glyco_hydro_19; 1.
 DR ProDom; PD000574; -; 1.
 DR PROSITE; PS00773; CHITINASE_19_1; 1.
 DR PROSITE; PS00774; CHITINASE_19_2; 1.
 DR Signal; Hydrolase; Glycosidase.
 FT SIGNAL 1
 FT CHAIN 26
 FT SEQUENCE 439 AA; 48496 MW; 684AAAE330FAE605 CRC64;

Query Match 57.8%; Score 52; DB 10; Length 439;
 Best Local Similarity 69.2%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DB 3 RRRPYLRPPRP 15
 50 RRRPYLRPPRP 62

RESULT 25
 Q39184 PRELIMINARY; PRT; 467 AA.
 AC Q39184;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PROTEIN KINASE.
 GN AME2 OR F4P12_270.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kuromori T., Yamamoto M.;
 RT "A. thaliana genes encoding protein kinases of a new family.";
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bloembergen H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
 RL Salanoubat M.;
 RN Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; D45354; BAA08215.1; -
 DR EMBL; AL132966; CAB67664.1; -
 DR HSSP; P24941; IA01.
 DR Mendel; 6951; Arch; 1349; 6951.
 DR InterPro: IPR000719; -
 DR InterPro: IPR002290; -
 DR Pfam; PF00069; pkinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50108; PROTEIN_KINASE_ST; 1.


```

RC STRAIN-BERKELEY;
RA MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale B., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu Q., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "the genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
DR EMBL: AE003826; AAC22283.1; -
DR InterPro: IPR002965; -
DR PRINTS: PRO1217; PRICHEXTNSN
SO SEQUENCE 94 AA; 10600 MW; F845534F483091D1 CRC64;

Query Match 56.7%; Score 51; DB 5; Length 94;
Best Local Similarity 62.5%; Pred. No. 4.9;
Matches 10; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

OY 4 PRPPYLR---PRPP 15
DB 51 PRPPHPPHPPHPP 66

RESULT 29
ID 090027 PRELIMINARY; PRT: 409 AA.
OS 090027
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHEICAL 45.5 KDA PROTEIN.
GN L5683.03.
OC Leishmania major.
OS Leishmania major.
RA Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.
OK NCBI_TaxID=5664;
RN SEQUENCE FROM N.A.
RP STRAIN-FRIEDLIN;
RC STRAIN-FRIEDLIN;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RX MEDLINE-98146435; PubMed-9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.,
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998)
DR EMBL: AL139794; CAC22630.1; -
KW Hypothetical protein.
SO SEQUENCE 409 AA; 45498 MW; DD09360FAE396E74 CRC64;

Query Match 56.7%; Score 51; DB 5; Length 409;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 RRRPPYLRPPRP 15
DB 69 RRRPPHPPHPPPP 83

RESULT 30
ID 09V4V1 PRELIMINARY; PRT: 417 AA.
AC 09V4V1
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE C614747 PROTEIN.
GN C614747.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phnydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chame M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```


RA ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Smith H.O.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003836; AAF59074.1; -
 DR FlyBase: FBgn0033306; CG14747.
 DR InterPro: IPR002965; -
 DR PRINTS: PR01217; PRICHEXTENS.
 SQ SEQUENCE 417 AA; 45701 MW; 376CC135371B1EC6 CRC64;

Query Match 56.7%; Score 51; DB 5; Length 417;
 Best Local Similarity 64.3%; Pred. No. 18;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRRPPYLRPPRP 15
 DB 308 RRRPPYLRPPRP 321

RESULT 31
 Q9V4J6 PRELIMINARY; PRT; 1307 AA.
 AC Q9V4J6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CG11101 PROTEIN.
 GN CG11101.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sultion G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Adill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Bertone D., Borkan M.R., Borman B.P., Bhandari D., Bolshakov S.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelira S., Fleischmann W.,
 RA Fostel C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkham I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003841; AAF59274.2; -
 DR HSSP: P35555; 1EMN.
 DR FlyBase: FBgn0033157; CG11101.
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001881; -
 DR InterPro: IPR002965; -
 DR Pfam: PF00008; EGF_2.
 DR PRINTS: PR01217; PRICHEXTENS.
 DR PROSITE: PS00010; ASX_HYDROXYL. 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS01187; EGF_CA; 1.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 1307 AA; 139890 MW; 42827878C1B77510 CRC64;

Query Match 56.7%; Score 51; DB 5; Length 1307;
 Best Local Similarity 64.3%; Pred. No. 51;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRRPPYLRPPRP 15
 DB 304 RRRPPYLRPPRP 317

RESULT 32
 Q9IBT9 PRELIMINARY; PRT; 3325 AA.
 AC Q9IBT9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE MAJOR TEGUMENT PROTEIN.
 GN UL36.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
 NCBI_TaxID=10390;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=GA.
 RX Jones D., Lee L., Lin J.L., Kung H.J., Tilletson J.R.;
 RA MEDLINE=92237304; PubMed=1315048;
 RA "Marek disease virus encodes a basic-leucine zipper gene resembling
 RT the fos/jun oncogenes that is highly expressed in lymphoblastoid
 RT tumors."
 RT Proc. Natl. Acad. Sci. U.S.A. 89:4042-4046(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GA;
 RA Lee L.F., Wu P., Sui D., Ren D., Kung H.J.;
 RT "The complete UL Sequence of Serotype I Marek's Disease Virus."
 RT Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2000).
 DR EMBL: AF147806; AAF66771.1; -
 DR InterPro: IPR002965; -
 DR PRINTS: PR01217; PRICHEXTENS.
 SQ SEQUENCE 3325 AA; 365698 MW; 5D4A84EF719B99FD CRC64;

Query Match 56.7%; Score 51; DB 14; Length 3325;
 Best Local Similarity 58.3%; Pred. No. 1,2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 RRRPPYLRPPRP 15
 DB 2745 RRRPPYLRPPRP 2756

RESULT 33

Q9FI95 ID Q9FI95 PRELIMINARY; PRT; 151 AA.
 AC Q9FI95;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K18J17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RC MEDLINE=99397451; PubMed=10470850;
 RA Kameko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen
 RT p1 and TAC clones.";
 RL DNA Res. 6:183-195(1999).
 DR EMBL: AB017060; BAB10796.1;
 SR SEQUENCE 151 AA; 17137 MW; D8BDEB53061EA32 CRC64;

Query Match 56.1%; Score 50.5; DB 10; Length 151;
 Best Local Similarity 60.0%; Pred. No. 8.7;
 Matches 9; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 2 RRPRLPP-PPRLPP 15
 :||:|:|:|:|:|:|
 Db 57 QRPQPLPLPQPP 71

RESULT 34
 ID 024102 PRELIMINARY; PRT; 249 AA.
 AC 024102;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE MTN4 (FRAGMENT).
 GN MTN4.
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Fabales; Fabaceae; Papilionoideae; Medicago.
 OX NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. JEMALONG J5; TISSUE=ROOT MODULE;
 RA Gamas P.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 GN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. JEMALONG J5; TISSUE=ROOT MODULE;
 RC MEDLINE=96212994; PubMed=8634476;
 RA Gamas P., de Carvalho Niebel F., Lescure N., Cullimore J.;
 RT "Use of a subtractive hybridization approach to identify new Medicago
 RT truncatula genes induced during root nodule development.";
 RL EMBL: Y15372; CAA75594.1; -;
 DR HSSP: P24337; IHVP.
 DR Mendel. 27080; Medtr.1531;27080.
 DR InterPro: IPR000528; -;
 DR InterPro: IPR003612; -;
 DR Pfam: PF00279; LTP; 1.
 DR SMART: SM00499; AAI; 1.
 FT NON_TER 1
 FT SEQUENCE 249 AA; 26923 MW; 4BF9256A0FDD1318 CRC64;

Query Match 56.1%; Score 50.5; DB 10; Length 249;
 Best Local Similarity 50.0%; Pred. No. 14;
 Matches 8; Conservative 4; Mismatches 1; Indels 3; Gaps 1;

QY 3 RRPRLPP-PPRLPP 15
 :||:|:|:|:|:|:|
 Db 12 QRPQPLPLPQPP 27

RESULT 35
 ID 09FP63 PRELIMINARY; PRT; 282 AA.
 AC 09FP63;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE P0475H04.3 PROTEIN.
 GN P0475H04.3.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0475H04.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002871; BAB20637.1; -;
 SR SEQUENCE 282 AA; 30036 MW; F3FC2A0C64D7B6E1 CRC64;

Query Match 56.1%; Score 50.5; DB 10; Length 282;
 Best Local Similarity 64.7%; Pred. No. 15;
 Matches 11; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 2 RRPRLPP-PPRLPP 15
 :||:|:|:|:|:|:|
 Db 39 RRPRLPP-PPRLPP 55

RESULT 36
 ID 088399 PRELIMINARY; PRT; 1479 AA.
 AC 088399; 088401; 088402; 088403; 088404; 035404;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE SYNAPTOJANIN 2 (EC 3.1.3.56) (SYNAPTIC INOSITOL-1,4,5-TRISPHOSPHATE 5-
 DE PHOSPHATASE 2).
 GN SYN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA TO ZETA).
 RC MEDLINE=98300294; PubMed=9636665;
 RX Seet L.-F., Cho S., Hessel A., Dumont D.J.;
 RA "Molecular cloning of multiple isoforms of synaptotjanin 2 and
 RT assignment of the gene to mouse chromosome 17A2.3.1.";
 RL Biochem. Biophys. Res. Commun. 247:116-122(1998).
 RN [2]
 RP SEQUENCE OF 43-1479 FROM N.A. (ISOFORMS 2A AND 2B).
 RC STRAIN=C57BL/6J;
 RX MEDLINE=98113197; PubMed=9442075;
 RA Khvotchev M., Sudhof T.C.;
 RT "Developmentally regulated alternative splicing in a novel
 RT synaptotjanin.";
 RL J. Biol. Chem. 273:2306-2311(1998).
 CC -I- FUNCTION: INOSITOL 5-PHOSPHATASE WHICH MAY BE INVOLVED IN DISTINCT

CC MEMBRANE TRAFFICKING AND SIGNAL TRANSDUCTION PATHWAYS.
 CC -1- CATALYTIC ACTIVITY: D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,4-BISPHOSPHATE + PHOSPHATE
 CC -1- SUBUNIT: THE ISOFORMS WITH PROLINE-RICH REGIONS BIND TO GMB2 AND
 CC ISOFORM 2A BINDS TO OMP25 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY ASSOCIATED WITH THE
 CC PARTICULATE FRACTIONS (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS: 2A, 2B, ALPHA, BETA, DELTA,
 CC EPSILON, GAMMA AND ZETA. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, WITH THE HIGHEST
 CC LEVELS IN HEART AND BRAIN. ISOFORM 2B IS EXPRESSED IN THE
 CC PERIPHERAL TISSUES WHILE ISOFORM 2A IS EXPRESSED IN THE BRAIN.
 CC -1- DEVELOPMENTAL STAGE: ALTERNATIVE SPLICING IS A DEVELOPMENTALLY
 CC LATE EVENT, OCCURRING ONLY AFTER THE FIRST POSTNATAL WEEK AFTER
 CC THE GENERATION OF NEURONS AND INITIAL SYNAPTOGENESIS.
 CC -1- DOMAIN: ISOFORMS WITH THE C-TERMINAL PROLINE-RICH REGION MEDIATE
 CC BINDING ONLY TO THE SH3 DOMAIN-CONTAINING PROTEIN GRB2.
 CC -1- DOMAIN: IN ISOFORMS 2A AND ZETA, AA 1237-1276 BINDS THE PD2 DOMAIN
 CC OF OMP25 (BY SIMILARITY).
 CC -1- SIMILARITY: IN THE CENTRAL SECTION: BELONGS TO THE INOSITOL-1,4,5-
 CC TRISPHOSPHATE 5-PHOSPHATASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SAC1 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- CAUTION: THE SEQUENCE SHOWN IS A MERGE BETWEEN BETA AND EPSILON
 CC ISOFORMS, FOR WHICH THE SEQUENCES ARE INCOMPLETE AT THE C-TERMINUS
 CC AND N-TERMINUS, RESPECTIVELY.
 CC EMBL: AF041857; AAC40141.1; -;
 CC EMBL: AF041859; AAC40143.1; -;
 CC EMBL: AF041858; AAC40142.1; -;
 CC EMBL: AF041860; AAC40144.1; -;
 CC EMBL: AF041861; AAC40145.1; -;
 CC EMBL: AF041862; AAC40146.1; -;
 CC EMBL: AF026123; AAC3137.1; -;
 CC MGD; MG1:1201671; SynJ2.
 CC InterPro: IPR000300; -;
 CC InterPro: IPR000504; -;
 CC InterPro: IPR002013; -;
 CC InterPro: IPR002965; -;
 CC Pfam: PF000076; trm; 1.
 CC Pfam: PF00783; IPec; 1.
 CC Pfam: PF02383; SyJa_N; 1.
 CC PRINTS: PR01217; PRICHEXTENSN.
 CC SMART: SM00128; IPec; 1.
 CC KJ HydroLase: Alternative splicing: Multigene family.
 CC EMBL: 1 449
 CC EMBL: 450 7
 CC EMBL: 889 968
 CC EMBL: 1109 1112
 CC EMBL: 1354 1357
 CC EMBL: 1 85
 CC EMBL: 508 508
 CC EMBL: 1068 1113
 CC EMBL: 1237 1276
 CC EMBL: 1237 1276
 CC EMBL: 1237 1276
 CC EMBL: 1237 1301
 CC EMBL: 1237 1304
 CC EMBL: 1277 1479
 CC EMBL: 1302 1479
 CC EMBL: 55 57
 CC EMBL: 68 70
 CC EMBL: 86 86

FT CONFLICT 101 101 A -> G (IN REF. 2).
 FT CONFLICT 104 104 T -> F (IN REF. 2).
 FT CONFLICT 119 119 L -> A (IN REF. 2).
 FT CONFLICT 153 153 G -> C (IN REF. 2).
 FT CONFLICT 179 179 N -> D (IN REF. 2).
 FT CONFLICT 238 238 T -> A (IN REF. 2).
 FT CONFLICT 263 263 R -> G (IN REF. 2).
 FT CONFLICT 278 278 G -> E (IN REF. 2).
 FT CONFLICT 297 297 K -> Q (IN REF. 2).
 FT CONFLICT 304 304 P -> L (IN REF. 2).
 FT CONFLICT 358 358 Q -> K (IN REF. 2).
 FT CONFLICT 362 362 E -> D (IN REF. 2).
 FT CONFLICT 420 420 I -> T (IN REF. 2).
 FT CONFLICT 437 437 T -> G (IN REF. 2).
 FT CONFLICT 534 534 V -> M (IN REF. 2).
 FT CONFLICT 533 533 A -> T (IN REF. 2).
 FT CONFLICT 576 576 S -> G (IN REF. 2).
 FT CONFLICT 580 580 V -> I (IN REF. 2).
 FT CONFLICT 583 583 I -> V (IN REF. 2).
 FT CONFLICT 674 674 L -> F (IN REF. 2).
 FT CONFLICT 682 682 V -> I (IN REF. 2).
 FT CONFLICT 773 773 AV -> TI (IN REF. 2).
 FT CONFLICT 831 832 PQ -> AN (IN REF. 2).
 FT CONFLICT 904 904 K -> R (IN REF. 2).
 FT CONFLICT 979 979 L -> I (IN REF. 2).
 FT CONFLICT 1004 1004 S -> T (IN REF. 2).
 FT CONFLICT 1013 1019 DYLEDE -> MFLKMK (IN AAC40144).
 FT CONFLICT 1017 1017 E -> D (IN REF. 2).
 FT CONFLICT 1023 1025 VDQ -> ADE (IN REF. 2).
 FT CONFLICT 1040 1040 N -> D (IN REF. 2).
 FT CONFLICT 1047 1048 SL -> AS (IN REF. 2).
 FT CONFLICT 1134 1134 Y -> C (IN REF. 2).
 FT CONFLICT 1185 1185 S -> G (IN REF. 2).
 FT CONFLICT 1196 1196 I -> T (IN REF. 2).
 FT CONFLICT 1209 1209 T -> S (IN REF. 2).
 FT CONFLICT 1214 1214 A -> V (IN REF. 2).
 FT CONFLICT 1220 1220 A -> V (IN REF. 2).
 FT CONFLICT 1243 1246 NEEP -> OCEO (IN REF. 2).
 FT CONFLICT 1269 1269 MISSING (IN REF. 2).
 FT CONFLICT 1280 1280 L -> F (IN REF. 2).
 FT CONFLICT 1284 1284 K -> R (IN REF. 2).
 FT CONFLICT 1288 1288 G -> R (IN REF. 2).
 FT CONFLICT 1292 1292 S -> D (IN REF. 2).
 FT CONFLICT 1299 1299 E -> D (IN REF. 2).
 FT CONFLICT 1302 1302 S -> P (IN REF. 2).
 FT CONFLICT 1311 1311 TVESP -> AVELTS (IN REF. 2).
 FT CONFLICT 1315 1315 Q -> P (IN REF. 2).
 FT CONFLICT 1343 1343 N -> S (IN REF. 2).
 FT CONFLICT 1347 1347 V -> L (IN REF. 2).
 FT CONFLICT 1359 1361 PSTL -> SPP (IN REF. 2).
 FT CONFLICT 1366 1372 HPLCLOV -> PLTYPOM (IN REF. 2).
 FT CONFLICT 1380 1380 R -> I (IN REF. 2).
 FT CONFLICT 1385 1385 H -> D (IN REF. 2).
 FT CONFLICT 1400 1400 N -> D (IN REF. 2).
 FT CONFLICT 1417 1417 N -> NN (IN REF. 2).
 FT CONFLICT 1429 1429 V -> L (IN REF. 2).
 FT CONFLICT 1433 1433 N -> T (IN REF. 2).
 FT CONFLICT 1440 1441 EP -> DS (IN REF. 2).
 FT CONFLICT 1448 1449 LA -> VV (IN REF. 2).
 FT CONFLICT 1462 1462 L -> F (IN REF. 2).
 FT CONFLICT 1466 1467 VT -> MA (IN REF. 2).
 FT CONFLICT 1479 AA; 163584 MW; 2261A6F0620DB2ED CRC64;

Query Match 56.1%; Score 50.5; DB 11; Length 1479;
 Best Local Similarity 71.4%; Pred. No. 66;
 Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 3 RRRPPLP-RRPP 15
 DB 1099 RRRPPLP-RRPP 1112

```

RESULT 37
Q9QTW0 PRELIMINARY; PRT; 96 AA.
AC 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE ORF2 PROTEIN (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBL_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYW MA993;
RA Mulyanto, Hijikata M., Matsushita M., Inokusumo G., Midjaya A.,
RA Sumarsidi D., Kanai K., Ohta Y., Mishiro S.;
RT "TT virus (TYV) genotypes in native and non-native prostitutes of
RT Irian Jaya, Indonesia: implication for non-occupational
RT transmission.";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031733; BAA86196.1;
FT NON_TER 1
SQ SEQUENCE 96 AA; 11160 MW; 12F657A821FB2871 CRC64;

Query Match 55.6%; Score 50; DB 14; Length 96;
Best Local Similarity 72.7%; Pred. No. 6.7;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPPVPRPP 15
DB 78 RPPVPRPP 88

RESULT 38
P74661 PRELIMINARY; PRT; 192 AA.
AC P74661;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DE 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE HYPOTHEICAL 20.9 KDA PROTEIN.
GN SUR160.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBL_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90917; BAA18779.1;
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 20865 MW; 3064186DC9776018 CRC64;

```

Query Match 55.6%; Score 50; DB 2; Length 192;
 Best Local Similarity 69.2%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRRPPYLRPP 14
 DB 3 RRRPPYLRPP 15

```

RESULT 39
O69260 PRELIMINARY; PRT; 344 AA.
AC 069260;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE LATENCY-RELATED OPEN READING FRAMES 1 AND 2, COMPLETE CDS.
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBL_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9105640; PubMed=2173765;
RA Kutish G., Mauprize T.H., Rock D.;
RT "Characterization of the latency-related transcriptionally active
RT region of the bovine herpesvirus 1 genome.";
RL J. Virol. 64:5730-5737(1990).
DR EMBL; M61143; AAA46060.1;
SQ SEQUENCE 344 AA; 32852 MW; D908839CB2082F66 CRC64;

```

Query Match 55.6%; Score 50; DB 14; Length 344;
 Best Local Similarity 64.3%; Pred. No. 21;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRRPPYLRPP 14
 DB 168 RRRPPYLRPP 181

```

RESULT 40
O9SVY1 PRELIMINARY; PRT; 383 AA.
AC O9SVY1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE ZINC FINGER-LIKE PROTEIN.
GN F15B8.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBL_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C., Quetier F., Salanoubat M.,
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049660; CAB41188.1;
DR InterPro; IPR000822;
DR Pfam; PF00096; zf-C2H2; 2
DR PROSITE; PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
DR SMART; SM00355; Znf_C2H2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 383 AA; 43337 MW; 9DDCE6E66A4506 CRC64;

```

Query Match 55.6%; Score 50; DB 10; Length 383;
 Best Local Similarity 64.3%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRRPPYLRPP 15
 DB 52 RRRPPYLRPP 65

Query Match	55.6%	Score 50	DB 10	Length 409
Best Local Similarity	69.2%	Pred. No. 24		
Matches	9	Conservative	0	Mismatches 4
				Indels 0
				Gaps 0
QY	3	RRPPPYLRRPP	15	
db	245	RRPPAPPPPPPP	257	

DR	PROSITE: PS00135; TRYPSIN. SER. 1.
DR	SMART: SM00020; Tryp_SPC. 1.
KW	Hydrolase; Serine protease; Signal.
FT	NOV TER
FT	SIGNAL
FT	CHAIN
SC	SEQUENCE

```
QY      2 RRRP--PYLPRRPP 15
Db      339 RPRRLQPEAPQPP 354
```

RESULT	43
09FF14	
ID	09FF14
AC	09FF14;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT	01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE	SIMILARITY TO UNKNOWN PROTEIN.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II;
OC	Brassicales; Brassicaceae; Arabidopsis.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-COLUMBIA;
RX	MEDLINE=97471969; PubMed=9330910;
RA	Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA	Miyajima N., Tabata S.;
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT	features of the 1.6 Mb regions covered by twenty physically assigned
RT	pl clones.";
RL	DNA Res. 4:215-230(1997).
DR	EMBL; AB005249; BAB09942.1; -
QO	SEQUENCE 464 AA; 52410 MW; 0673BFCA89AEB3B CRC64;

Query Match	55.6%;	Score 50;	DB 10;	Length 464;
Best Local Similarity	61.9%;	Pred. No. 27;		
Matches 13; Conservative	0;	Mismatches 2;	Indels 6;	Gaps 2

```
QY      1 RRR----PRPPYLRP--RPP 15
          .  |||  | || |||  |||
Db      52 RRRAPPPPPPLPRPCSRPP 72
```

RESULT	44	
Q9MA04	ID	PRELIMINARY;
Q9MA04	PRT:	593 AA.
AC	Q9MA04.	
DT	01-OCT-2000	(TREMBLrel. 15, Created)
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)
DE	F20B17.16.	
OS	Arabidopsis thaliana	(Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;	
OC	Brassicales; Brassicaceae; Arabidopsi.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP-	SEQUENCE FROM N.A.	
RA	Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,	
RA	Shinn P., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,	
RA	Conway A., Gonzales A., Hansen N., Hong B., Koo T., Lam B., Lee J.,	
RA	Leuz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,	

RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayer A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome
 RT 1.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altif H., Bel B., Chin C., Chlou J., Choi E.,
 RA Com L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thayer A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC010793; AAF6118.1; -;
 DR InterPro: IPR002965; -;
 DR PRINTS: PR01217; PRICEXTENS.
 DR SEQUENCE 593 AA; 68056 MW; 3E188CF6A2F8E61A CRC64;

Query Match 55.6%; Score 50; DB 10; Length 593;
 Best Local Similarity 72.7%; Pred. No. 34;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 RPPYLRPRPP 15
 |||||
 Db 5 RPPYLRPRPP 15

RESULT 45
 ID Q9VNA6 PRELIMINARY; PRT; 596 AA.
 AC Q9VNA6;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CG16708 PROTEIN.
 GN CG16708.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Ayday A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes W., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA - Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Munkulov G., Misha N.V., Mody B., Mody C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003603; AAF52040.1; -;
 DR FLYbase: FBgn0037315; CG16708.
 DR InterPro: IPR001206; -;
 DR Pfam: PF00781; DAGKC; 1;
 DR SMART: SM00046; DAGKC; 1;
 DR SEQUENCE 596 AA; 65622 MW; 91410AC18AAC4660 CRC64;

Query Match 55.6%; Score 50; DB 5; Length 596;
 Best Local Similarity 72.7%; Pred. No. 34;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 RPPYLRPRPP 15
 |||||
 Db 294 RPPYLRPRPP 304

Search completed: September 24, 2001, 10:07:41
 Job time: 141 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 24, 2001, 10:08:35 : Search time 44.37 Seconds
(without alignments)
15.030 Million cell updates/sec

Title: US-09-276-868-4
Perfect score: 64
Sequence: 1 RRRPRPPYLP R 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSR/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSR/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSR/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSR/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSR/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSR/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSR/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSR/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSR/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSR/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSR/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSR/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSR/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSR/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSR/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSR/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSR/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSR/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSR/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSR/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSR/gcgdata/geneseq/AA2000.DAT:*
22: /SIDSR/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	64	100.0	11	21	AAB26886	PR-39 derived anol
2	64	100.0	14	17	AAW01450	Leukocyte O2- prod
3	64	100.0	14	19	AAW5725	Proline/Arginine r
4	64	100.0	15	21	AAB26885	PR-39 derived ang1
5	64	100.0	19	17	AAW01452	Leukocyte O2- prod
6	64	100.0	26	17	AAW01447	Leukocyte O2- prod
7	64	100.0	26	19	AAW5723	Proline/Arginine r
8	64	100.0	39	14	AAW30491	Antibacterial pept
9	64	100.0	39	17	AAW01446	Leukocyte O2- prod
10	64	100.0	39	17	AAW94446	Synducin peptide (
11	64	100.0	39	17	AAW9121	Magainin-derived a

12	64	100.0	39	19	AAW5722	Proline/Arginine r
13	64	100.0	39	22	AAB26888	PR-39 peptide used
14	64	100.0	44	21	AAW51194	E. coli AMP gene p
15	49	76.6	23	17	AAW01451	Leukocyte O2- prod
16	48	75.0	8	21	AAB26887	PR-39 derived ang1
17	47	73.4	18	16	AAR79211	Bactenein peptide
18	47	73.4	20	19	AAW75730	Proline/Arginine r
19	47	73.4	23	16	AAR79209	Bactenein peptide
20	47	73.4	35	16	AAR79212	Bactenein peptide
21	47	73.4	59	17	AAR94448	Synducin peptide (
22	47	73.4	59	19	AAW66400	Cationic peptide B
23	47	73.4	59	22	AAW51197	E. coli AMP gene B
24	47	73.4	62	22	AAW51197	Human prostate can
25	45	70.3	272	21	AAW52125	Human alpha 1-6 fu
26	44	68.8	575	18	AAW22124	Pig alpha 1-6 fuco
27	44	68.8	153	21	AAW54649	Zea mays protein f
28	43	67.2	153	21	AAW54649	Amino acid sequenc
29	41	64.1	106	20	AAV29263	HSV-2 strain SB5 C
30	41	64.1	121	19	AAW72052	Parapox virus prot
31	41	64.1	548	18	AAW40306	Dog TSL0q23.3 gene
32	40.5	63.3	36	20	AAV07466	Dog TSL0q23.3 gene
33	40.5	63.3	430	20	AAV07458	Streptococcus anti
34	40	62.5	273	21	AAV92838	S. avermitilis HPP
35	40	62.5	306	21	AAV51614	Arabidopsis thalia
36	40	62.5	309	21	AAV51614	Arabidopsis thalia
37	40	62.5	309	21	AAV51614	Arabidopsis thalia
38	40	62.5	324	21	AAV51614	Arabidopsis thalia
39	40	62.5	324	21	AAV51614	Arabidopsis thalia
40	40	62.5	333	21	AAV51614	Arabidopsis thalia
41	40	62.5	333	21	AAV51614	Arabidopsis thalia
42	40	62.5	421	16	AAW74172	Tyrosinase and tyr
43	40	62.5	426	16	AAW74172	Tyrosinase and tyr
44	40	62.5	478	16	AAW74174	Chloroplast transi
45	40	62.5	506	21	AAW20574	Perilla frutescens

ALIGNMENTS

RESULT 1
AAB26886 standard; peptide; 11 AA.
AC AAB26886;
DT 01-FEB-2001 (first entry)
DE PR-39 derived angiogenesis regulatory peptide 2.
XX Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
KW myocardial ischaemia; proteasome.
XX Synthetic.
OS WO200057895-A1.
XX WO200057895-A1.
XX 05-OCT-2000.
XX 16-MAR-2000; 2000MO-US07050.
XX 26-MAR-1999; 99US-0276868.
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX Simons M, Gao Y;
XX WPI; 2000-628319/60.
XX Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
XX infarction, by administering a PR-39 oligopeptide that regulates
XX enzymatic activity of proteasomes
XX Claim 13; Page 41; 51pp; English.
PS

XX This invention relates to a method for the stimulation of angiogenesis in
CC situ within a targeted collection of viable cells. The method comprises
CC introducing, into the cytoplasm, at least 1 member of the PR-39
CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
CC Part of the proteolytic activity of the proteasomes is selectively
CC altered so as to stimulate angiogenesis. The method is used to induce
CC angiogenesis in tissue that has suffered anoxia or infarction,
CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
CC study the mechanisms that control angiogenesis. The present sequence
CC represents a PR-39 derived peptide which interacts with the proteasome
CC and can be used in the method of the invention.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 21; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPPYLPR 11
| | | | | | | | | | |
Db 1 rrrpppylpr 11

RESULT 2
ID AAW01450 standard; peptide; 14 AA.
XX
AC AAW01450;
XX
DT 18-JUN-1997 (first entry)
XX

Leukocyte O2- production inhibitor peptide PR14.

XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
XX antimicrobial peptide; small intestine; human; neutrophil; bacteria;
XX DNA synthesis; protein synthesis; inhibitor; syndecan expression;
XX mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
XX tissue damage; oxygen radical; inflammatory disease; therapy.
XX
OS Synthetic.
XX
PN W09632129-A1.
XX
PD 17-OCT-1996.
XX
PF 10-APR-1996; 96WO-US04674.
XX
PR 10-APR-1995; 95US-0419066.
XX
PA (UNIV) UNIV KANSAS STATE RES FOUND.
XX
PI Blecha F, Shi J;
XX
XX WPI; 1996-476842/47.

XX Inhibition of leukocyte superoxide anion prodn. and attraction of
XX leukocytes - using peptide(s) partic. based on antimicrobial PR-39
XX
XX Claim 12; Page 28; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
XX antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
XX isolated from porcine small intestine, and has also been identified in
XX human and porcine neutrophils. PR39 kills bacteria by interfering with
XX DNA and/or protein synthesis. PR39 also induces syndecan expression on
XX mesenchymal cells. Syndecans are important in wound repair, showing that
XX PR39 can be used in wound repair, as well as in antibacterial agents.
XX These sequences, and PR39, can be used in the method of the invention.
XX The method of the invention is for inhibiting leukocyte superoxide anion
XX (O2-) production. The method comprises administering to a leukocyte a
XX peptide (such as this sequence) capable of inhibiting leukocyte O2-

CC production. The peptides can be used as medicaments for fighting
CC infection by attracting leukocytes to a wound site and restricting
CC tissue damage at the wound site caused by excessive oxygen radicals
CC produced by these leukocytes. They can also be used to develop products
CC for treating inflammatory disease states.
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 64; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPPYLPR 11
| | | | | | | | | | |
Db 1 rrrpppylpr 11

RESULT 3
ID AAW75725 standard; peptide; 14 AA.
XX
AC AAW75725;
XX
DT 19-NOV-1998 (first entry)
XX
DE Proline/Arginine rich peptide.
XX
KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
KW coronary bypass; organ transplantation surgery.
XX
OS Synthetic.
XX
PN W09833690-A1.
XX
PD 20-AUG-1998.
XX
PF 17-FEB-1998; 98WO-US03207.
XX
PR 16-FEB-1998; 98US-0024975.
XX
PR 18-FEB-1997; 97US-0802306.
XX
PA (UNIV) UNIV KANSAS STATE RES FOUND.
XX
PI Blecha F, Ross CR, Shi J;
XX
XX WPI; 1998-495359/42.

XX Reduction of reperfusion injury in temporarily occluded blood
XX vessels - by administration of a peptide which is rich in proline
XX or arginine residues
XX
XX Claim 3; Page 15; 35pp; English.

XX Sequences AAW75722-W75732 are proline/arginine rich peptides that upon
XX administration into a mammal's bloodstream reduce reperfusion injury
XX (production of reactive oxygen species, neutrophil adherence to
XX endothelium, and extravasation of neutrophils). These peptides have two
XX requirements: they contain the consensus sequence PXXP, where P is a
XX proline residue and X is any amino acid residue, which has been found to
XX inhibit superoxide production, and secondly they have arginine residues
XX adjacent to these motifs, required for effective inhibition. It was
XX established by structural and function analysis that a peptide should
XX ideally contain 4 or 6 of these motifs, and that inhibitory activity is
XX correlated with the increase of length of peptides. The effectiveness
XX of these peptides was determined by investigating the production of the
XX neutrophil superoxide anion, and also the inhibition of neutrophil
XX chemotaxis. From this, it was found that all of the peptides inhibited
XX NADPH oxidase to some extent. All of the peptides also inhibit
XX neutrophil oxidase activity. PR-39 is believed, to be the most potent
XX endogenous down regulator of NADPH oxidase yet discovered, and from the
XX data produced, it can be suggested to be involved in eliminating or

CC reducing the reperfusion injury induced adhesion and extraction of
 CC neutrophils. The peptides are also useful in connection with surgical
 CC procedures such as coronary bypass and organ transplantation surgery.

XX Sequence 14 AA;

Query Match 100.0%; Score 64; DB 19; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.0059;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 |||||
 Db 1 rrrpppylpr 11

RESULT 4
 AAB26885
 ID AAB26885 standard; peptide; 15 AA.

XX AAB26885;

CT 01-FEB-2001 (first entry)

XX PR-39 derived angiogenesis regulatory peptide 1.

XX Angiogenesis: stimulation; PR-39; anoxia; myocardial infarction;
 KW myocardial ischemia; proteasome.

XX Synthetic.

XX WO200057895-A1.

XX 05-OCT-2000.

XX 16-MAR-2000; 2000WO-US07050.

XX 26-MAR-1999; 99US-0276868.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Simons M, Gao Y;

XX WPI; 2000-628319/60.

XX Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PT infarction, by administering a PR-39 oligopeptide that regulates
 PT enzymatic activity of proteasomes

XX Claim 12; Page 40; 51pp; English.

XX This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.

CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents a PR-39 derived peptide which interacts with the proteasome
 CC and can be used in the method of the invention.

XX Sequence 15 AA;

Query Match 100.0%; Score 64; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 |||||
 Db 1 rrrpppylpr 11

RESULT 5
 AAW01452
 ID AAW01452 standard; peptide; 19 AA.

XX AAW01452;

XX 18-JUN-1997 (first entry)

XX Leukocyte O2- production inhibitor peptide PR19.

KW Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.

XX Synthetic.

XX WO9632129-A1.

XX 17-OCT-1996.

XX 10-APR-1996; 96WO-US04674.

XX 10-APR-1995; 95US-0419066.

XX (UNIV) UNIV KANSAS STATE RES FOUND.

XX Blecha F, Shi J;

XX WPI; 1996-476842/47.

XX Inhibition of leukocyte superoxide anion prodn. and attraction of
 PT leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 XX Disclosure; Page 27; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
 CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
 CC isolated from porcine small intestine, and has also been identified in
 CC human and porcine neutrophils. PR39 kills bacteria by interfering with
 CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
 CC mesenchymal cells. Syndecans are important in wound repair, showing that
 CC PR39 can be used in wound repair, as well as in antibacterial agents.
 CC These sequences, and PR39, can be used in the method of the invention.
 CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O2-) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.

XX Sequence 19 AA;

Query Match 100.0%; Score 64; DB 17; Length 19;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 |||||
 Db 1 rrrpppylpr 11

RESULT 6
 AAW01447
 ID AAW01447 standard; peptide; 26 AA.

XX AAW01447;

XX 18-JUN-1997 (first entry)
 DT Leukocyte O2- production inhibitor peptide PR26.
 XX
 DE Inhibitor: Leukocyte O2- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 XX DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.
 OS Synthetic.
 XX WO9632129-A1.
 PN
 XX 17-OCT-1996.
 PD
 XX 10-APR-1996; 96WO-US04674.
 PF
 XX 10-APR-1995; 95US-0419066.
 PR
 XX (UNIV) UNIV KANSAS STATE RES FOUND.
 PA
 XX Blecha F, Shi J;
 PI
 XX WPI: 1996-476842/47.
 DR
 XX Inhibition of leukocyte superoxide anion produ. and attraction of
 PT leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 PS
 XX Claim 3; Page 26; 45pp; English.
 CC AAM01447-W01454 represent fragments of the proline-arginine rich
 CC antimicrobial peptide PR39 (see AAM01446). The PR39 sequence was first
 CC isolated from porcine small intestine, and has also been identified in
 CC human and porcine neutrophils. PR39 kills bacteria by interfering with
 CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
 CC mesenchymal cells. Syndecans are important in wound repair, showing that
 CC PR39 can be used in wound repair, as well as in antibacterial agents.
 CC These sequences, and PR39, can be used in the method of the invention.
 CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O2-) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.
 CC
 SO Sequence 26 AA:
 XX
 XX
 Query Match 100.0%; Score 64; DB 17; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPYLPR 11
 |||||||||
 DB 1 rrrprpylpr 11
 RESULT 7
 ID AAM75723 standard; peptide; 26 AA.
 XX
 AC AAM75723;
 XX
 DT 19-NOV-1998 (first entry)
 DE Proline/Arginine rich peptide PR-26.
 KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
 KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
 XX

KM coronary bypass; organ transplantation surgery.
 XX
 OS Synthetic.
 XX WO9835690-A1.
 PN
 XX 20-AUG-1998.
 PD
 XX 17-FEB-1998; 98WO-US03207.
 PF
 XX 16-FEB-1998; 98US-0024975.
 PR
 XX 18-FEB-1997; 97US-0802306.
 PR
 XX (UNIV) UNIV KANSAS STATE RES FOUND.
 PA
 XX Blecha F, Ross CR, Shi J;
 PI
 XX WPI: 1998-495359/42.
 DR
 XX Reduction of reperfusion injury in temporarily occluded blood
 PT vessels - by administration of a peptide which is rich in proline
 PT or arginine residues
 PS
 XX Claim 3; Page 14-15; 35pp; English.
 CC Sequences AAM75722-W75732 are proline/arginine rich peptides that upon
 CC administration into a mammal's bloodstream reduce reperfusion injury
 CC (production of reactive oxygen species, neutrophil adherence to
 CC endothelium, and extravasation of neutrophils). These peptides have two
 CC requirements: they contain the consensus sequence PXXP, where P is a
 CC proline residue and X is any amino acid residue, which has been found to
 CC inhibit superoxide production, and secondly they have arginine residues
 CC adjacent to these motifs, required for effective inhibition. It was
 CC established by structural and function analysis that a peptide should
 CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
 CC correlated with the increase of length of peptides. The effectiveness
 CC of these peptides was determined by investigating the production of the
 CC neutrophil superoxide anion, and also the inhibition of neutrophil
 CC chemotaxis. From this, it was found that all of the peptides inhibited
 CC NADPH oxidase to some extent. All of the peptides also inhibit
 CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
 CC endogenous down regulator of NADPH oxidase yet discovered, and from the
 CC data produced, it can be suggested to be involved in eliminating or
 CC reducing the reperfusion injury induced adhesion and extraction of
 CC neutrophils. The peptides are also useful in connection with surgical
 CC procedures such as coronary bypass and organ transplantation surgery.
 CC
 SO Sequence 26 AA:
 XX
 XX
 Query Match 100.0%; Score 64; DB 19; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPYLPR 11
 |||||||||
 DB 1 rrrprpylpr 11
 RESULT 8
 ID AAR30491 standard; peptide; 39 AA.
 XX
 AC AAR30491;
 XX
 DT 12-MAY-1993 (first entry)
 DE Antibacterial peptide.
 KW Pig; small intestine; endocrine; gram negative; bacteria; therapeutic;
 KW veterinary medicine; prophylactic.
 XX
 OS Sus scrofa domestica.

XX WO9222578-A.
 XX 23-DEC-1992.
 XX 10-JUN-1992; 92WO-SE00394.
 XX 14-JUN-1991; 91SE-0001838.
 XX (BOMA/) BOMAN H G.
 XX (JOER/) JOERNVALL H.
 XX (LEED/) LEE J.
 XX (MUTR/) MUTR V.
 XX Boman HG, Joernvall H, Lee J, Mutt V;
 XX WPI; 1993-018080/02.
 XX
 XX New anti-bacterial polypeptide - active against Gram negative
 XX bacteria
 XX Claim 1; Page 10; 15pp; English.
 XX
 XX This peptide was isolated from the small intestine of a pig. The
 XX small intestine is an important endocrine organ and many of its
 XX physiologically active peptides have been isolated from it. This
 XX peptide inhibits the growth of, and may kill, bacteria, pref. gram
 XX negative bacteria. This peptide or its functional derivatives may be
 XX used in human or veterinary medicine for therapeutic or prophylactic
 XX use.
 XX
 XX Sequence 39 AA:

Query Match 100.0%; Score 64; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 |||||
 DB 1 rrrpppylpr 11

RESULT 9

AAW01446
 ID AAW01446 standard; peptide; 39 AA.

AC AAW01446;

DT 18-JUN-1997 (first entry)

DE Leukocyte O2- production inhibitor peptide PR39.

KW Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.

OS Synthetic.

PN WO9632129-A1.

PD 17-OCT-1996.

PF 10-APR-1996; 96WO-US04674.

PR 10-APR-1995; 95US-0419066.

PA (UNIV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Shi J;

XX

DR WPI; 1996-476842/47.
 XX Inhibition of leukocyte superoxide anion prodn. and attraction of
 XX leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 XX Claim 2; Page 26; 45pp; English.
 XX
 XX This sequence represents the proline-arginine rich antimicrobial peptide
 XX PR39. The PR39 sequence was first isolated from porcine small intestine,
 XX and has also been identified in human and porcine neutrophils. PR39
 XX kills bacteria by interfering with DNA and/or protein synthesis. PR39
 XX also induces syndecan expression on mesenchymal cells. Syndecans are
 XX important in wound repair, showing that PR39 can be used in wound repair,
 XX as well as in antibacterial agents. This sequence, and the fragments of
 XX it shown in AAW01447-W01454, can be used in the method of the invention.
 XX The method of the invention is for inhibiting leukocyte superoxide anion
 XX (O2-) production. The method comprises administering to a leukocyte a
 XX peptide (such as this sequence) capable of inhibiting leukocyte O2-
 XX production. The peptides can be used as medicaments for fighting
 XX infection by attracting leukocytes to a wound site and restricting tissue
 XX damage at the wound site caused by excessive oxygen radicals produced by
 XX these leukocytes. They can also be used to develop products for treating
 XX inflammatory disease states.
 XX
 XX Sequence 39 AA:

Query Match 100.0%; Score 64; DB 17; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 |||||
 DB 1 rrrpppylpr 11

RESULT 10

AAW01446
 ID AAW01446 standard; peptide; 39 AA.

AC AAW01446;

DT 05-NOV-1996 (first entry)

DE Synducin peptide (PR-39) induces syndecan expression.

KW Synducin; induction; expression; syndecan-1; syndecan-4; surface;
 KW mesenchymal cell; fibroblast; epithelial; PR-39; treatment; stasis;
 KW decubitus; ulcers; keloids; skin burns; ischemic tissues;
 KW hypercoagulation states; prevention; tumour metastasis; restenosis;
 KW inhibition; angiogenesis; proliferation; endothelial.

OS Synthetic.

PN WO9609322-A2.

PD 28-MAR-1996.

PF 22-SEP-1995; 95WO-US12080.

PR 22-SEP-1994; 94US-0310722.

PA (CHIL-) CHILDRENS MEDICAL CENT.

PI Bernfield M, Gallo RL;

DR WPI; 1996-188401/19.

PT Modulating mesenchymal interaction by administration of synducin
 PT used in the treatment of wounds, tumours, restenosis, etc

PS Claim 4; Page 26; 34pp; English.

XX

CC The present peptide is a syndecin, which induces the expression of
CC syndecan-1 and syndecan-4 on the surface of mesenchymal cells, esp.
CC fibroblasts and epithelial cells. The 36 N-terminal amino acids of
CC the peptide were found to be identical to the 36 N-terminal amino
CC acids of PR-39, a Pro and Arg rich antibacterial peptide previously
CC found in porcine intestine (W09222578). Syndecins may be used in
CC the treatment of stasis and decubitus ulcers, keloids, skin burns,
CC ischemic tissues and hypercoagulation states, prevention of tumour
CC metastasis, restenosis inhibition and endothelial cell angiogenesis
CC and proliferation induction.
CC Human microvascular endothelial cells were assayed for syndecan-4
CC expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
CC the present peptide (10 microm) or a blank, to give respective
CC cell surface syndecan-4 values (MOD/m in) of approx. 1.75, 1.70,
CC 1.80 and 0.95.

CC Sequence 39 AA;

Query Match 100.0%; Score 64; DB 17; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
Db 1 rrrpppylpr 11

RESULT 11

AAR99121
ID AAR99121 standard; peptide; 39 AA.

AC AAR99121;

DT 28-OCT-1996 (first entry)

DE Magainin-derived antimicrobial STD-inhibiting peptide, MSI-1312.

KW STD; sexually transmitted disease; HIV; human immunodeficiency virus;
KW herpes simplex virus; HSV; Neisseria gonorrhoeae; Candida; Chlamydia;
KW Magainin; antimicrobial; squalamine.

OS Synthetic.

Key Location/Qualifiers
FT Modified-site 39 /note="amdated"

PN W09608270-A2.

PD 21-MAR-1996.

PF 13-SEP-1995; 95WO-US11675.

PR 13-SEP-1994; 94US-0305475.

PA (MAGA-) MAGAININ PHARM INC.

PI Bedi G, Jacob L, Williams T, Zasloff M;

WPI; 1996-179725/18.

PT Inhibiting sexually transmitted disease e.g. HIV or herpes simplex -
PT by administering magainin antimicrobial or squalamine cpd. to
PT inhibit transmission

PS Example 1; Page 32; 60pp; English.

CC AAR99116-R99123 are antimicrobial, magainin-analogue peptides that may
CC be used to treat sexually transmitted diseases (STDs) caused by
CC Chlamydia, HIV, herpes simplex virus, Neisseria gonorrhoeae or
CC Candida infection. The peptides inhibit STDs by either killing the
CC infectious organism, impeding the infection mechanism or

CC interrupting the replication cycle of the organism. Squalamine (an
CC antihistrol host defence molecule of the dog fish shark Squalus
CC acanthias) and Pgla (a frog antimicrobial peptide) analogues may
CC also be useful in inhibiting STD infection and transmission.

CC Sequence 39 AA;

Query Match 100.0%; Score 64; DB 17; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
Db 1 rrrpppylpr 11

RESULT 12

AAW75722
ID AAW75722 standard; peptide; 39 AA.

AC AAW75722;

DT 19-NOV-1998 (first entry)

DE Proline/Arginine rich peptide PR-39.

KW Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
KW coronary bypass; organ transplantation surgery.

OS Synthetic.

PN W09835690-A1.

PD 20-AUG-1998.

PF 17-FEB-1998; 98WO-US03207.

PR 16-FEB-1998; 98US-0024975.

PR 18-FEB-1997; 97US-0802306.

PA (UNITV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Ross CR, Shi J;

WPI; 1998-495359/42.

PT Reduction of reperfusion injury in temporarily occluded blood
PT vessels - by administration of a peptide which is rich in proline
PT or arginine residues

PS Claim 3; Page 14; 35pp; English.

CC Sequences AAW75722-W75732 are proline/arginine rich peptides that upon
CC administration into a mammal's bloodstream reduce reperfusion injury
CC (production of reactive oxygen species, neutrophil adherence to
CC endothelium, and extravasation of neutrophils). These peptides have two
CC requirements: they contain the consensus sequence PXXP, where P is a
CC proline residue and X is any amino acid residue, which has been found to
CC inhibit superoxide production, and secondly they have arginine residues
CC adjacent to these motifs, required for effective inhibition. It was
CC established by structural and function analysis that a peptide should
CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
CC correlated with the increase of length of peptides. The effectiveness
CC of these peptides was determined by investigating the production of the
CC neutrophil superoxide anion, and also the inhibition of neutrophil
CC chemotaxis. From this, it was found that all of the peptides inhibited
CC NADPH oxidase to some extent. All of the peptides also inhibit
CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
CC endogenous down regulator of NADPH oxidase yet discovered, and from the
CC data produced, it can be suggested to be involved in eliminating or
CC reducing the reperfusion injury induced adhesion and extraction of

CC neutrophils. The peptides are also useful in connection with surgical
 CC procedures such as coronary bypass and organ transplantation surgery.

XX
 SQ Sequence 39 AA;

Query Match 100.0%; Score 64; DB 19; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 RRRPRPYLPR 11
 |||||
 1 rrrprpylpr 11

RESULT 13

AAB26888 standard; peptide; 39 AA.

XX AAB26888;

XX 01-FEB-2001 (first entry)

DE PR-39 peptide used in angiogenesis control.

XX Angiogenesis: stimulation; PR-39; anoxia; myocardial infarction;
 KW myocardial ischemia; proteasome.

XX Synthetic.

XX WO200057895-A1.

XX 05-OCT-2000.

XX 16-MAR-2000; 2000WO-US07050.

XX 26-MAR-1999; 99US-0276868.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Simons M, Gao Y;

XX WPI; 2000-628319/60.

PT Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PT infarction, by administering a PR-39 oligopeptide that regulates
 PT enzymatic activity of proteasomes -

PS Disclosure; Page 21; 51pp; English.

XX This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
 CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents the PR-39 peptide from which peptide used in the method of
 CC the invention are derived.

XX Sequence 39 AA;

Query Match 100.0%; Score 64; DB 21; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPYLPR 11
 |||||
 1 rrrprpylpr 11

RESULT 14
 ID AAB51194 standard; Peptide; 44 AA.

XX AAB51194;

XX 22-MAR-2001 (first entry)

DE E. coli AMP gene PR39 amino acid sequence.

XX Escherichia coli; E. coli; AMP gene; anti-microbial peptide;

KW screening; preservation; food; feed; paint formulation; detergent;

KW cosmetic; medical device; prosthetic implant; disinfectant;

XX microbial infection; tumour.

XX Escherichia coli.

XX WO200073433-A1.

XX 07-DEC-2000.

XX 29-MAY-2000; 2000WO-DK00287.

XX 31-MAY-1999; 99DK-0000766.

XX (NOVO) NOVO NORDISK AS.

XX Kristensen H;

XX WPI; 2001-070965/08.

PT Screening for nucleotide sequences encoding antimicrobial peptides by
 PT generating peptide libraries in microorganisms, inducing expression of
 PT peptides, selecting cells and recovering nucleotide sequences from
 PT cells -

PS Example 1; Page 34; 59pp; English.

XX The present invention describes a plasmid which is ligated with a pool
 CC of nucleotide sequences (NT) linked to an inducible promoter, to express
 CC a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
 CC optionally linked to a signal (P)). transformed into host cells and
 CC cultured in presence of an inducer to induce expression of the NT. A
 CC method of screening (I) a pool of nucleotide sequences to select a
 CC nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
 CC with the pool of NT; (b) transforming host cells which are sensitive to
 CC the peptide with the ligated plasmids; (c) screening the transformed
 CC cells to select viable cells; (d) cultivating the viable cells in the
 CC presence of an inducer to induce expression of NT; (e) selecting cells
 CC according to the effect of the inducer on cell growth; and (f) recovering
 CC NT encoding the peptide from the selected cells. (I) is useful for
 CC screening a pool of nucleotide sequences to select a nucleotide sequence
 CC encoding a peptide which is an antimicrobial peptide or an antimicrobial
 CC enzyme active on bacteria and for finding and preparing a composition for
 CC treatment of human or animal. The antimicrobial peptide obtained using
 CC (I) may be employed in preservation of e.g. food/feed, paint
 CC formulations, detergents, cosmetics, medical devices such as prosthetic
 CC implants and also to disinfect and/or kill microbial cells on an object
 CC e.g. as an disinfectant for the treatment of biofilm. The peptides are
 CC useful for treating microbial infections and/or tumours. Peptides with
 CC improved bio-activity can be developed using (I). The peptides have no
 CC negative effect on normal mammalian and/or eukaryotic cells. The present
 CC sequence represents an AMP (antimicrobial peptide) gene amino acid
 CC sequence, which is used in an example from the present invention.

XX Sequence 44 AA;

Query Match 100.0%; Score 64; DB 22; Length 44;
 Best Local Similarity 100.0%; Pred. No. 0.016;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPYLP 11
 |||||
 Db 3 rrrppylpr 13

RESULT 15

AAW01451
 ID AAW01451 standard; peptide: 23 AA.
 AC AAW01451;
 XX
 DT 18-JUN-1997 (first entry)
 DE Leukocyte O2- production inhibitor peptide PR23.
 XX
 KW Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KW DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.
 XX
 OS Synthetic.
 PN WO9632129-A1.
 PD 17-OCT-1996.
 XX
 PF 10-APR-1996; 96WO-US04674.
 XX
 PR 10-APR-1995; 95US-0419066.
 XX
 PA (UNITV) UNIV KANSAS STATE RES FOUND.
 XX
 PI Blecha F, Shi J;
 DR WPI: 1996-476842/47.
 XX
 PT Inhibition of leukocyte superoxide anion prodn. and attraction of
 PR leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 PS
 XX Disclosure: Page 27; 45pp; English.
 CC AAW01447-W01454 represent fragments of the proline-arginine rich
 CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
 CC isolated from porcine small intestine, and has also been identified in
 CC human and porcine neutrophils. PR39 kills bacteria by interfering with
 CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
 CC mesenchymal cells. Syndecans are important in wound repair, showing that
 CC PR39 can be used in wound repair, as well as in antibacterial agents.
 CC These sequences, and PR39, can be used in the method of the invention.
 CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O2-) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.
 CC
 SO Sequence 23 AA;

Query Match 76.6%; Score 49; DB 17; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PPPPYLP 11
 |||||
 Db 1 ppppylpr 8

RESULT 16
 AAB26887

ID AAB26887 standard; peptide: 8 AA.

AC AAB26887;

DT 01-FEB-2001 (first entry)

DE PR-39 derived angiogenesis regulatory peptide 3.

KW Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
 KW myocardial ischaemia; proteasome.

OS Synthetic.

PN WO200057895-A1.

PD 05-OCT-2000.

PF 16-MAR-2000; 2000WO-US07050.

PR 26-MAR-1999; 99US-0276868.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Simons M, Gao Y;

DR WPI: 2000-628319/60.

PT Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PR infarction, by administering a PR-39 oligopeptide that regulates
 XX enzymatic activity of proteasomes

PS Claim 14; Page 41; 51pp; English.

CC This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
 CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents a PR-39 derived peptide which interacts with the proteasome
 CC and can be used in the method of the invention.
 CC
 SO Sequence 8 AA;

Query Match 75.0%; Score 48; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPYLP 8
 |||||
 Db 1 rrrppylpr 8

RESULT 17

AA79211
 ID AAR79211 standard; Peptide: 18 AA.

AC AAR79211;

DT 05-MAR-1996 (first entry)

DE Bactenecin peptide deriv. C, with antibacterial activity.

KW Bac5; Bac7; bactenecin; antibacterial; food preservative.

OS Synthetic.

PN EP665239-A1.

XX

PO 02-AUG-1995.
XX
XX 26-JAN-1994; 94EP-0200176.
XX
XX 26-JAN-1994; 94EP-0200176.
XX
XX (CNDR) CONSIGLIO NAZ DELLE RICERCHE.
PA (UTR-) UNIV TRIESTE.
XX
PI Gennaro R, Romeo D, Scocchi M, Skerlavaj B;
XX
XX WPI; 1995-264826/35.
XX
XX New peptide derivs. of bactenecin - useful as antibacterial agents,
PT with an activity comparable to Bac5 and Bac7
XX
XX Claim 4; Page 5; 7pp; English.
XX
XX AAR79209-12 have antibacterial activity which is at least comparable to
CC that of the known bactenecins, Bac5 and Bac7, in molar terms. They may
CC be used as food preservatives. The peptides may be prepd. using Fmoc
CC techniques on an automated continuous flow synthesiser.
XX
SQ Sequence 18 AA;

Query Match 73.4%; Score 47; DB 16; Length 18;
Best Local Similarity 81.8%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRRPPYLPYR 11
| | | | | | | |
DB 2 RIRPPRPYLPYR 12

RESULT 18
AAR75730
ID AAM75730 standard; peptide: 20 AA.
XX
XX AAM75730;
XX
XX 19-NOV-1998 (first entry)
XX
XX Proline/Arginine rich peptide Bac7.
XX
XX Proline; arginine; peptide; reperfusion injury; neutrophil; endothelium;
KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
KW coronary bypass; organ transplantation surgery.
XX
XX Synthetic.
XX
XX WO9835690-A1.
XX
XX 20-AUG-1998.
XX
XX 17-FEB-1998; 98WO-US03207.
XX
XX 16-FEB-1998; 98US-0024975.
XX
XX 18-FEB-1997; 97US-0802306.
XX
XX (UNIV) UNIV KANSAS STATE RES FOUND.
XX
XX Blecha F, Ross CR, Shi J;
XX
XX WPI; 1998-495359/42.
XX
XX Reduction of reperfusion injury in temporarily occluded blood
PT vessels - by administration of a peptide which is rich in proline
PT or arginine residues
XX
XX Claim 3; Page 18; 35pp; English.
XX
XX Sequences AAM75722-W75732 are proline/arginine rich peptides that upon

CC administration into a mammal's bloodstream reduce reperfusion injury
CC (production of reactive oxygen species, neutrophil adherence to
CC endothelium, and extravasation of neutrophils). These peptides have two
CC requirements: they contain the consensus sequence PXXP, where P is a
CC proline residue and X is any amino acid residue, which has been found to
CC inhibit superoxide production, and secondly they have arginine residues
CC adjacent to these motifs, required for effective inhibition. It was
CC established by structural and function analysis that a peptide should
CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
CC correlated with the increase of peptides. The effectiveness
CC of these peptides was determined by investigating the production of the
CC neutrophil superoxide anion, and also the inhibition of neutrophil
CC chemotaxis. From this, it was found that all of the peptides inhibited
CC NADPH oxidase to some extent. All of the peptides also inhibit
CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
CC endogenous down regulator of NADPH oxidase yet discovered, and from the
CC data produced, it can be suggested to be involved in eliminating or
CC reducing the reperfusion injury induced adhesion and extraction of
CC neutrophils. The peptides are also useful in connection with surgical
CC procedures such as coronary bypass and organ transplantation surgery.
XX
SQ Sequence 20 AA;

Query Match 73.4%; Score 47; DB 19; Length 20;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRRPPYLPYR 11
| | | | | | | |
DB 2 RIRPPRPYLPYR 12

RESULT 19
AAR79209
ID AAR79209 standard; Peptide: 23 AA.
XX
XX AAR79209;
XX
XX 05-MAR-1996 (first entry)
XX
XX Bactenecin peptide deriv. A, with antibacterial activity.
XX
XX Bac5; Bac7; bactenecin; antibacterial; food preservative.
XX
XX Synthetic.
XX
XX EP665239-A1.
XX
XX 02-AUG-1995.
XX
XX 26-JAN-1994; 94EP-0200176.
XX
XX 26-JAN-1994; 94EP-0200176.
XX
XX (CNDR) CONSIGLIO NAZ DELLE RICERCHE.
XX
XX (UTR-) UNIV TRIESTE.
XX
XX Gennaro R, Romeo D, Scocchi M, Skerlavaj B;
XX
XX WPI; 1995-264826/35.
XX
XX New peptide derivs. of bactenecin - useful as antibacterial agents,
PT with an activity comparable to Bac5 and Bac7
XX
XX Claim 2; Page 5; 7pp; English.
XX
XX AAR79209-12 have antibacterial activity which is at least comparable to
CC that of the known bactenecins, Bac5 and Bac7, in molar terms. They may
CC be used as food preservatives. The peptides may be prepd. using Fmoc
CC techniques on an automated continuous flow synthesiser.
XX
XX Sequence 23 AA;

Query Match 73.4%; Score 47; DB 16; Length 23;
 Best Local Similarity 81.8%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 1 | | | | |
 2 rlrpprprlpr 12

Db

RESULT 20
 AAR79212
 ID AAR79212 standard; Peptide; 35 AA.
 AC AAR79212;
 DT 05-MAR-1996 (first entry)
 DE Bacteneclin peptide deriv. D, with antibacterial activity.
 DE Bac5; Bac7; bacteneclin; antibacterial; food preservative.
 OS Synthetic.
 PN EP665239-A1.
 PD 02-AUG-1995.
 PE 26-JAN-1994; 94EP-0200176.
 PR 26-JAN-1994; 94EP-0200176.
 PA (CMDR) CONSIGLIO NAZ DELLE RICERCHE.
 PA (UYTR-) UNIV TRIESTE.
 PI Genaro R, Romeo D, Scocchi M, Skerlavaj B;
 DR WPI: 1995-264826/35.
 PS New peptide derivs. of bacteneclin - useful as antibacterial agents,
 PT with an activity comparable to Bac5 and Bac7
 PS Claim 5; Page 5; 7pp; English.
 CC AAR79209-12 have antibacterial activity which is at least comparable to
 CC that of the known bacteneclins, Bac5 and Bac7, in molar terms. They may
 CC be used as food preservatives. The peptides may be prep'd. using Fmoc
 CC techniques on an automated continuous flow synthesiser.
 SQ Sequence 35 AA;

Query Match 73.4%; Score 47; DB 16; Length 35;
 Best Local Similarity 81.8%; Pred. No. 2.8;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 1 | | | | |
 2 rlrpprprlpr 12

Db

RESULT 21
 AAR94448
 ID AAR94448 standard; peptide; 59 AA.
 AC AAR94448;
 DT 05-NOV-1996 (first entry)
 DE Synducin peptide (Bac-7) induces syndecan expression.
 DE Synducin; induction; expression; syndecan-1; syndecan-4; surface;

KW mesenchymal cell; fibroblast; epithelial; Bac-7; treatment; stasis;
 KW decubitus; ulcers; keloids; skin burns; ischemic tissues;
 KW hypercoagulation states; prevention; tumour metastasis; restenosis;
 KW inhibition; angiogenesis; proliferation; endothelial.
 OS Synthetic.
 PN 1.13
 PD Peptide
 PE /note=
 PR 1.18
 PA /note=
 PI 1.23
 PT Peptide
 PS /note=
 CC 5.23
 CC /note=
 CC 15.42
 CC /note=
 CC "peptide fragment giving 0 % increase in
 CC syndecan-1 expression on NIH 3T3 cells"
 CC "peptide fragment giving 10 % increase
 CC in syndecan-1 expression on NIH 3T3
 CC cells"
 CC "peptide fragment giving 230 % increase
 CC in syndecan-1 expression on NIH 3T3
 CC cells"
 CC "peptide fragment giving 250 % increase
 CC in syndecan-1 expression on NIH 3T3
 CC cells"
 CC "peptide fragment giving 230 % increase
 CC in syndecan-1 expression on NIH 3T3
 CC cells"
 CC "peptide fragment giving 10 % increase
 CC in syndecan-1 expression on NIH 3T3
 CC cells"
 CC "peptide fragment giving 0% increase in
 CC syndecan-1 expression on NIH 3T3 cells"

W09609322-A2.
 28-MAR-1996.
 22-SEP-1995; 95WO-US12080.
 22-SEP-1994; 94US-0310722.
 (CHIL-) CHILDRENS MEDICAL CENT.
 Bernfield M, Gallo RL;
 WPI: 1996-188401/19.
 DR Modulating mesenchymal interaction by administration of synducin -
 PT used in the treatment of wounds, tumours, restenosis, etc
 PS Claim 4; Page 27; 34pp; English.
 CC The present peptide (previously known as an antibacterial peptide),
 CC specifically the claimed peptide comprising residues 1-18, is a
 CC synducin, which induces the expression of syndecan-1 and syndecan-4
 CC on the surface of mesenchymal cells, esp. fibroblasts
 CC and epithelial cells. Synducins may be used in the treatment of
 CC stasis and decubitus ulcers, keloids, skin burns, ischemic tissues
 CC and hypercoagulation states, prevention of tumour metastasis,
 CC restenosis inhibition and endothelial cell angiogenesis and
 CC proliferation induction.
 CC Human microvascular endothelial cells were assayed for syndecan-4
 CC expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
 CC the present peptide (10 microM) or a blank, to give respective
 CC cell surface syndecan-4 values (MOD/m in) of approx. 1.75, 1.70,
 CC 2.75 and 0.95.
 SQ Sequence 59 AA;

Query Match 73.4%; Score 47; DB 17; Length 59;
 Best Local Similarity 81.8%; Pred. No. 4.5;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
 1 | | | | |
 2 rlrpprprlpr 12

Db

XX (NOVO) NOVO NORDISK AS.
 PA
 XX
 XX

PI Kristensen H;
 XX
 XX

DR WPI: 2001-070965/08.
 XX
 XX

Screening for nucleotide sequences encoding antimicrobial peptides by
 generating peptide libraries in microorganisms, inducing expression of
 peptides, selecting cells and recovering nucleotide sequences from
 cells

Example 1; Page 34; 59pp; English.

The present invention describes a plasmid which is ligated with a pool
 of nucleotide sequences (NT) linked to an inducible promoter, to express
 a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
 optionally linked to a signal (P)), transformed into host cells and
 cultured in presence of an inducer to induce expression of the NT. A
 method of screening (I) a pool of nucleotide sequences to select a
 nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
 with the pool of NT; (b) transforming host cells which are sensitive to
 the peptide with the ligated plasmids; (c) screening the transformed
 cells to select viable cells; (d) cultivating the viable cells in the
 presence of an inducer to induce expression of NT; (e) selecting cells
 according to the effect of the inducer on cell growth; and (f) recovering
 NT encoding the peptide from the selected cells. (I) is useful for
 screening a pool of nucleotide sequences to select a nucleotide sequence
 encoding a peptide which is an antimicrobial peptide or an antimicrobial
 enzyme active on bacteria and for finding and preparing a composition for
 treatment of human or animal. The antimicrobial peptide obtained using
 (I) may be employed in preservation of e.g. food/feed, paint
 formulations, detergents, cosmetics, medical devices such as prosthetic
 implants and also to disinfect and/or kill microbial cells on an object
 e.g. as an disinfectant for the treatment of biofilm. The peptides are
 useful for treating microbial infections and/or tumours. Peptides with
 improved bio-activity can be developed using (I). The peptides have no
 negative effect on normal mammalian and/or eukaryotic cells. The present
 sequence represents an AMP (antimicrobial peptide) gene amino acid
 sequence, which is used in an example from the present invention.

Sequence 62 AA;

Query Match 73.4%; Score 47; DB 22; Length 62;
 Best Local Similarity 81.8%; Pred. No. 4.7;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLP 11
 | | | | | | |
 DB 4 RTRPPPYLP 14

RESULT 25

AAB57119
 ID AAB57119 standard; Protein: 272 AA.

AC AAB57119;

XX 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1697.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX
 XX

OS Homo sapiens.

XX MO200055174-A1.

XX 21-SEP-2000.
 PD
 XX

PF 08-MAR-2000; 2000WO-US05988.
 XX
 XX

PR 12-MAR-1999; 99US-0124270.
 XX
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 (ROSE/) ROSEN C A.
 XX

PI Rosen CA, Ruben SM;
 XX

XX WPI: 2000-587513/55.
 DR N-PSDB: AAF16322.

Prostate cancer associated gene sequences, referred to as prostate
 cancer antigens, useful for treatment, prevention, and diagnosis of
 disorders such as prostate cancer -

Claim 11; Page 2169-2170; 2338pp; English.

AAFI5566 to AAF16505 encode the human prostate cancer associated
 proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytosolic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAF57303 represent sequences used in the exemplification of the present
 CC invention.
 XX

Sequence 272 AA;

Query Match 70.3%; Score 45; DB 21; Length 272;
 Best Local Similarity 80.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLP 10
 | | | | | | |
 DB 222 RTRPPPYLP 231

RESULT 26
 AAW22125
 ID AAW22125 standard; Protein: 575 AA.

AC AAW22125;

XX 05-MAR-1998 (first entry)

DE Human alpha 1-6 fucosyltransferase.

XX Alpha 1-6 fucosyltransferase; enzyme; pig; human; fucose transfer;
 KW guanosine diphosphate; sugar chain synthesis; modification; antibody;
 KW GICNAc; cancer diagnosis.
 KW
 XX

OS Homo sapiens.

XX MO9727303-A1.

XX 31-JUL-1997.

PF 23-JAN-1997; 97WO-JP00171.

PR 22-JUL-1996; 96JP-0192260.

PR 24-JAN-1996; 96JP-0010365.

PR 21-JUN-1996; 96JP-0161648.

PR 24-JUN-1996; 96JP-0162813.
 XX (TOYM) TOYO BOSEKI KK.
 PA Shiba T, Taniguchi N, Uozumi N, Yanagidani S;
 PI MPI: 1997-393690/36.
 DR N-PSDB; AAT76574.
 XX
 XX Human or pig alpha 1-6 fucosyl:transferase and DNA encoding it - for
 PT synthesis and modification of sugar chains and used as an antigen
 PT for production of diagnostic antibodies
 PS Claim 17; Page 39-43; 61pp; Japanese.
 XX
 CC AAW22124 and AAW22125 represent the pig and human alpha 1-6
 CC fucosyltransferases of the invention, respectively. The enzyme transfers
 CC fucose from guanosine diphosphate to the 6-hydroxyl group of the GlcNAc
 CC nearest to R in the receptor molecule: (GlcNAc-beta 1-2Man-alpha 1-6)
 CC (GlcNAc-beta 1-2Man-alpha 1-3)Man-beta 1-4GlcNAc-R to give
 CC 1-4GlcNAc-beta 1-4(Fuc-alpha 1-6)GlcNAc-R. It has an optimum pH of about
 CC 7.0 (pig) or 7.5 (human), and is stable over the pH range 4-10 after 5
 CC hours at 4 degrees C. The optimum working temperature of the
 CC alpha 1-6 fucosyltransferase is 30-37 degrees C. A bivalent metal is
 CC not required for activity of the enzyme, and the enzyme is not inhibited
 CC in the presence of 5 mM EDTA. The enzyme is useful in the synthesis and
 CC modification of sugar chains, and as antigen for the production of
 CC antibodies recognising the enzyme. The antibodies can be used for the
 CC diagnosis of cancer and other diseases.
 CC
 CC Sequence 575 AA:
 XX

Query Match 68.8%; Score 44; DB 18; Length 575;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PRPYLP 10
 |||||
 Db 299 PRPYLP 305

RESULT 27
 AAW22124
 ID AAW22124 standard; Protein; 575 AA.
 XX
 AC AAW22124;
 XX
 DT 05-MAR-1998 (first entry)
 XX
 DE Pig alpha 1-6 fucosyltransferase.
 XX
 KW Alpha 1-6 fucosyltransferase; enzyme; pig; human; fucose transfer;
 KW guanosine diphosphate; sugar chain synthesis; modification; antibody;
 KW GlcNAc; cancer diagnosis.
 XX
 OS Sus scrofa.
 XX
 PN WO9727303-A1.
 XX
 PD 31-JUL-1997.
 XX
 PF 23-JAN-1997; 97WO-JP00171.
 XX
 XX 22-JUL-1996; 96JP-0192260.
 XX 24-JAN-1996; 96JP-0010365.
 PR 21-JUN-1996; 96JP-0161648.
 PR 24-JUN-1996; 96JP-0162813.
 XX
 PA (TOYM) TOYO BOSEKI KK.
 PA Shiba T, Taniguchi N, Uozumi N, Yanagidani S;
 PI

XX MPI: 1997-393690/36.
 DR N-PSDB; AAT76573.
 XX
 XX Human or pig alpha 1-6 fucosyl:transferase and DNA encoding it - for
 PT synthesis and modification of sugar chains and used as an antigen
 PT for production of diagnostic antibodies
 PS Claim 4; Page 30-34; 61pp; Japanese.
 XX
 CC AAW22124 and AAW22125 represent the pig and human alpha 1-6
 CC fucosyltransferases of the invention, respectively. The enzyme transfers
 CC fucose from guanosine diphosphate to the 6-hydroxyl group of the GlcNAc
 CC nearest to R in the receptor molecule: (GlcNAc-beta 1-2Man-alpha 1-6)
 CC (GlcNAc-beta 1-2Man-alpha 1-3)Man-beta 1-4GlcNAc-R to give
 CC 1-4GlcNAc-beta 1-4(Fuc-alpha 1-6)GlcNAc-R. It has an optimum pH of about
 CC 7.0 (pig) or 7.5 (human), and is stable over the pH range 4-10 after 5
 CC hours at 4 degrees C. The optimum working temperature of the
 CC alpha 1-6 fucosyltransferase is 30-37 degrees C. A bivalent metal is
 CC not required for activity of the enzyme, and the enzyme is not inhibited
 CC in the presence of 5 mM EDTA. The enzyme is useful in the synthesis and
 CC modification of sugar chains, and as antigen for the production of
 CC antibodies recognising the enzyme. The antibodies can be used for the
 CC diagnosis of cancer and other diseases.
 CC
 CC Sequence 575 AA:
 XX

Query Match 68.8%; Score 44; DB 18; Length 575;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PRPYLP 10
 |||||
 Db 299 PRPYLP 305

RESULT 28
 AAG54649
 ID AAG54649 standard; Protein; 153 AA.
 XX
 AC AAG54649;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Zea mays protein fragment SEQ ID NO: 69704.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.
 XX
 OS Zea mays subsp. mays.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129645.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.

PR	23-APR-1999;	98US-0130891.
PR	28-APR-1999;	98US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132863.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136031.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137528.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138887.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	17-JUN-1999;	99US-0139453.
PR	18-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139465.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	02-JUL-1999;	99US-0142154.
PR	06-JUL-1999;	99US-0142055.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	98US-0130891.
PR	28-APR-1999;	98US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132863.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136031.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137528.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138887.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR</		

FR 14-OCT-1999; 99US-0159638.
 FR 18-OCT-1999; 99US-0159584.
 FR 21-OCT-1999; 99US-0160741.
 FR 21-OCT-1999; 99US-0160767.
 FR 21-OCT-1999; 99US-0160768.
 FR 21-OCT-1999; 99US-0160770.
 FR 21-OCT-1999; 99US-0160814.
 FR 21-OCT-1999; 99US-0160815.
 FR 22-OCT-1999; 99US-0160980.
 FR 22-OCT-1999; 99US-0160981.
 FR 22-OCT-1999; 99US-0160989.
 FR 25-OCT-1999; 99US-0161404.
 FR 25-OCT-1999; 99US-0161405.
 FR 25-OCT-1999; 99US-0161406.
 FR 26-OCT-1999; 99US-0161359.
 FR 26-OCT-1999; 99US-0161360.
 FR 26-OCT-1999; 99US-0161361.
 FR 28-OCT-1999; 99US-0161920.
 FR 28-OCT-1999; 99US-0161922.
 FR 28-OCT-1999; 99US-0161993.
 FR 29-OCT-1999; 99US-0162142.

Query Match 67.2%; Score 43; DB 21; Length 153;
 Best Local Similarity 70.0%; Pred. No. 37;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPPYLP R 11
 ||| ||| |||
 Db 16 rrrppfrrp 25

RESULT 29

AAI29263
 ID AAI29263 standard; Protein; 106 AA.

AC AAY29263;

DT 25-OCT-1999 (first entry)

DE Amino acid sequence of a virulence factor encoded by ORF41598.

KW Human pathogen; virulence polypeptide; virulence factor;

KW pathogenic infection; Pseudomonas aeruginosa infection.

OS Pseudomonas aeruginosa.

PN WO9927129-A1.

PD 03-JUN-1999.

XX 25-NOV-1998; 98WO-US25247.

PR 25-NOV-1997; 97US-0066517.

PA (GEO) GEN HOSPITAL CORP.

PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;

PI Ralme LG, Tan M, Tsongalis J;

DR WPI; 1999-357851/30.

XX Virulence factors useful in developing disease treatments

XX Disclosure; Fig 4; 228pp; English.

CC The present sequence represents a Pseudomonas aeruginosa polypeptide
 CC sequence. P. aeruginosa is an opportunistic human pathogen present in
 CC soil water and plants. The specification describes virulence polypeptides
 CC and nucleic acid sequence encoding such polypeptides. These sequences
 CC can be used to identify a compound which is capable of decreasing the
 CC expression of a pathogenic virulence factor. Compounds that inhibit
 CC the expression or activity of virulence factor polypeptides can be
 CC used to treat pathogenic infections, especially where the infection

CC is a P. aeruginosa infection.
 CC note: the sequences given in the specification were poorly legible, and
 CC in some instances assumptions were made as to the identity of the
 CC residue; it is therefore possible that the sequence given below is
 CC not entirely correct.

SO Sequence 106 AA;

Query Match 64.1%; Score 41; DB 20; Length 106;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPP R 7
 |||||
 Db 27 rrrppp 33

RESULT 30

AAW72052
 ID AAW72052 standard; Protein; 121 AA.

AC AAW72052;

DT 07-DEC-1998 (first entry)

DE HSV-2 strain SB5 Contig ID 50 ORF#1 protein.

KW HSV-2 strain SB5; immunological response induction; therapy;

KW antiviral identification; viral protein inhibitor.

OS Herpes simplex virus type 2.

FH Key Location/Qualifiers

FT Misc-difference 120..121

FT "note" encoded by CC"

PN WO9820016-A1.

PD 14-MAY-1998.

XX 31-OCT-1997; 97WO-US20016.

PR 09-JUN-1997; 97US-0049018.

PR 04-NOV-1996; 96US-0030279.

PA (SMK) SMITHKLINE BEECHAM CORP.

PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;

PI Esser KM, Leary JI;

DR WPI; 1998-286847/25.

DR N-PSDB; AAV62137.

PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention

PT and treatment of infection or inducing immunological response in

PT mammal

PS Claim 10; Page 60; 748pp; English.

CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 50.
 CC Based on homology, this sequence is a neurovirulence factor protein.
 CC The proteins can be used for the treatment or prevention of disease, to
 CC induce an immunological response in a mammal or to identify inhibitors,
 CC activators or novel antivirals. Antagonists of the proteins can be used
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
 CC it can also be used to induce an immunological response in a mammal.
 CC Sequence 121 AA;

Query Match 64.1%; Score 41; DB 19; Length 121;
 Best Local Similarity 69.2%; Pred. No. 57;
 Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;
 OY 1 RRRPRP--PYLPR 11
 ||||| | : ||
 Db 12 rrrprpgrptvpr 24

RESULT 31
 AAM40306
 ID AAM40306 standard; Protein; 548 AA.
 XX
 AC AAM40306;
 XX
 DT 07-JUL-1998 (first entry)
 XX
 DE Parapox virus protein kinase.
 XX
 KM Recombinant parapox virus; PPV; vaccine; diagnostic; protection;
 KM non-pathogen-specific immunostimulation; response; protein kinase;
 KW vascular endothelial growth factor; VEGF.
 XX
 OS Parapox virus.
 XX
 PN DEL19639601-A1.
 XX
 PD 04-SEP-1997.
 XX
 PF 26-SEP-1996; 96DE-1039601.
 XX
 PR 28-FEB-1996; 96DE-1007458.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Buettner M, Rziha H, Schmeer N, Strube W;
 XX
 DR WPI; 1997-436729/41.
 XX
 N-PSDB; AAT76874.
 PT Parapox virus strain D1701 HindIII fragment I, related recombinant
 PT viruses and expression plasmids - useful in vaccines, for expression
 PT of foreign DNA and for diagnosis
 XX
 PS Claim 28; Page -; 48pp; German.
 XX
 CC This sequence is protein kinase derived from a HindIII fragment from
 CC Parapox virus (PPV) strain D1701. Recombinant PPV, which may contain
 CC foreign DNA, are used in vaccines, to protect against the source of the
 CC foreign DNA, to produce a non-pathogen-specific immunostimulation or for
 CC expression of foreign DNA. The plasmids can be used to express
 CC PPV-specific gene sequences and in diagnostic reagents. DNA sequences
 CC that can be inserted (preferably 0.1-15 kb) are typically from herpes
 CC viruses, foot and mouth disease virus, influenza virus, Pasteurella,
 CC Salmonella, Chlamydia, Toxoplasma, Dirofilaria etc. The VEGF promoter can
 CC be used to express DNA in PPV. PPV are naturally immunostimulatory so
 CC vaccines based on them provide both a quickly developed non-specific
 CC protection and a long-lasting pathogen specific response.
 CC This sequence is not represented in the specification but is claimed.
 XX
 SQ Sequence 548 AA;

Query Match 64.1%; Score 41; DB 18; Length 548;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RRRPRP 7
 |||||
 Db 523 rrrprp 529
 RESULT 32

AAV07466
 ID AAV07466 standard; Protein; 36 AA.
 XX
 AC AAV07466;
 XX
 DT 16-JUL-1999 (first entry)
 XX
 DE Dog TS10q23.3 gene bases 1-108 translation product.
 XX
 KM Mutation; mutant; tumour suppressor; TS10q23.3; diagnosis; breast; ovary;
 KM Cowden's Syndrome; thyroid gland; endometrium; cancer; antibody; ELISA;
 KW enzyme linked immunosorbant assay; gene expression; dog.
 XX
 OS Canis familiaris.
 XX
 PN WO9910537-A1.
 XX
 PD 04-MAR-1999.
 XX
 PF 26-AUG-1998; 98WO-US17636.
 XX
 PR 30-APR-1998; 98US-0083563.
 PR 26-AUG-1997; 97US-0057750.
 XX
 PA (MYRI-) MYRIAD GENETICS INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Jasser SA, Pershouse MA, Steck P, Tavtigian SV;
 PI Yung WKA;
 XX
 DR WPI; 1999-190638/16.
 XX
 PT Newly isolated DNA or RNA polynucleotide encoding a mutant tumour
 PT suppressor (TS10q23.3) - useful for the diagnosis of Cowden's
 PT Syndrome and susceptibility to breast cancer
 XX
 PS Disclosure; Fig 9; 24pp; English.
 XX
 CC The invention relates to mutant genes encoding the tumour suppressor
 CC TS10q23.3. The TS10q23.3 mutant gene is useful for the diagnosis of
 CC Cowden's Syndrome in cells selected from the breast, ovaries, thyroid
 CC gland and endometrium (claimed). The mutant gene is also useful for
 CC diagnosing a subject who has a predisposition to breast cancer. Both
 CC methods involve antibodies, which specifically bind to a TS10q23.3,
 CC used in an ELISA assay to evaluate the level of TS10q23.3 expression.
 XX
 SQ Sequence 36 AA;

Query Match 63.3%; Score 40.5; DB 20; Length 36;
 Best Local Similarity 81.8%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 OY 1 RRRP-RPPYLP 10
 |||| | | |
 Db 13 rrrprppllp 23

RESULT 33
 AAV07458
 ID AAV07458 standard; Protein; 430 AA.
 XX
 AC AAV07458;
 XX
 DT 16-JUL-1999 (first entry)
 XX
 DE Dog TS10q23.3 gene bases 1-1290 translation product.
 XX
 KM Mutation; mutant; tumour suppressor; TS10q23.3; diagnosis; breast; ovary;
 KM Cowden's Syndrome; thyroid gland; endometrium; cancer; antibody; ELISA;
 KW enzyme linked immunosorbant assay; gene expression; dog.
 XX
 OS Canis familiaris.

XX MO9910537-A1.
 XX 04-MAR-1999.
 XX 26-AUG-1998; 98WO-US17636.
 XX 30-APR-1998; 98US-0083563.
 XX 26-AUG-1997; 97US-0057750.
 XX (MTRI-) MYRIAD GENETICS INC.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Jasser SA, Pershouse MA, Steck P, Tavligian SV,
 XX Yung WKA;
 XX WPI: 1999-190638/16.
 XX Newly isolated DNA or RNA polynucleotide encoding a mutant tumour
 XX suppressor (TS10q23.3) - useful for the diagnosis of Cowden's
 XX Syndrome and susceptibility to breast cancer
 XX
 XX Disclosure: Fig 9; 244pp; English.
 XX The invention relates to mutant genes encoding the tumour suppressor
 XX TS10q23.3. The TS10q23.3 mutant gene is useful for the diagnosis of
 XX Cowden's Syndrome in cells selected from the breast, ovaries, thyroid
 XX gland and endometrium (claimed). The mutant gene is also useful for
 XX diagnosing a subject who has a predisposition to breast cancer. Both
 XX methods involve antibodies, which specifically bind to a TS10q23.3,
 XX used in an ELISA assay to evaluate the level of TS10q23.3 expression.

DR WPI: 2000-350737/30.
 DR N-PSDB; AAA28414.
 XX
 XX New nucleic acid comprising Streptomyces tyrosinase and ORF-438 useful
 XX for treating disorders of tyrosinase expression and melanin
 XX biosynthesis
 XX
 XX Example 2; Fig 2B; 27pp; English.
 XX This is the Streptococcus antibioticus tyrosinase protein encoded on the
 XX bicistronic plmelsN nucleotide sequence insert comprising the
 XX S. antibioticus ORF-438 coding sequence coupled to the tyrosinase coding
 XX sequence via an IRES (internal ribosome entry site). In order to enhance
 XX expression of the bacterial gene in mammalian cells, the ORF-438 and
 XX tyrosinase gene TGA termination codons were altered to TAA. The Kozak
 XX consensus sequence was added upstream, immediately preceding the ATG
 XX initiation codon in each case to facilitate translation efficiency. The
 XX ORF-438 protein regulates copper incorporation, which is essential for
 XX the expression and function of tyrosinase and melanin production in
 XX Streptomyces. The invention concerns a bicistronic nucleotide sequence
 XX comprising a modified Streptomyces mel locus, such as the present
 XX sequence or Tyrosinase-IRES-ORF-438. Tyrosinase is the key enzyme for
 XX melanin biosynthesis and disorders of tyrosinase activity include
 XX Parkinson's disease, vitiligo and albinism. The bicistronic sequences
 XX are therefore useful in treatment of such disorders. Transfected
 XX mammalian cells are useful in determining the effect of a compound on
 XX pigmentation comprising culturing the cells in the presence or absence
 XX of the test compound and assessing the increase/decrease in melanin
 XX production. Additionally, the invention concerns inhibiting alopecia by
 XX expressing the cell cycle inhibitor gene p21 in hair follicle cells.

Query Match 63.3%; Score 40.5; DB 20; Length 430;
 Best Local Similarity 81.8%; Pred. No. 2.1e+02;
 Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 RRRP-RPYLP 10
 ||| ||| ||
 DB 13 rrrprppllp 23

Query Match 62.5%; Score 40; DB 21; Length 273;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRP-RPYLP 10
 ||| ||| ||
 DB 228 rthpspylp 237

RESULT 34
 AAY92838
 ID AAY92838 standard; Protein; 273 AA.

XX AC AAY92838;

XX 29-AUG-2000 (first entry)

DE Streptococcus antibioticus tyrosinase.

XX Bicistronic; plmelsN-ORF-438; tyrosinase; internal ribosome entry site;
 XX IRES; copper incorporation; melanin biosynthesis; Parkinson's disease;
 XX vitellogenin; albinism; alopecia; cell cycle inhibitor; p21.

XX Streptococcus antibioticus.

XX WO200024895-A2.

XX 04-MAY-2000.

XX 27-OCT-1999; 99WO-US25118.

XX 27-OCT-1998; 98US-0105725.

XX 27-OCT-1998; 98US-0105831.

XX (AMTI-) ANTICANCER INC.
 XX Zhao M;

RESULT 35
 AAY51614
 ID AAY51614 standard; Protein; 306 AA.

XX AC AAY51614;

XX 26-MAY-2000 (first entry)

DE S. avermitilis HPPD protein.

XX DOXS; 1-deoxy-D-xylulose-5-phosphate synthase; HPPD; GGPPOR; plant;
 XX p-hydroxyphenylpyruvate dioxygenase; tocopherol; vitamin K; chlorophyll;
 XX geranylgeranyl pyrophosphate oxidoreductase; carotenoid; transgenic.

XX Streptomyces avermitilis.

XX WO200008169-A1.

XX 17-FEB-2000.

XX 30-JUL-1999; 99WO-EP05467.

XX 05-AUG-1998; 98DE-1035219.

XX 01-OCT-1998; 98DE-1045216.

XX 01-OCT-1998; 98DE-1045224.

XX 01-OCT-1998; 98DE-1045231.

XX (SDNG-) SUNGENE GMBH & CO KGAA.
 XX Relndl A, Mejia PL, Palmas JME, Gracia MAC, Edneth M, Herbers K;

XX DR WPI; 2000-195578/17.
DR N-PSDB; AA288978.
XX
XX
PT Use of DNA encoding 1-deoxy-D-xylulose-5-phosphate synthase to produce
PT plants with increased tocopherol, vitamin K, chlorophyll and carotenoid
PT content -
XX
XX Example 10; Page 80-82; 94pp; German.
XX
CC This invention describes the novel use of a DNA sequence encoding
CC 1-deoxy-D-xylulose-5-phosphate synthase (DOXS), and optionally
CC p-hydroxyphenylpyruvate dioxygenase (HPPD) and/or geranylgeranyl-
CC pyrophosphate oxidoreductase (GGPOR), to produce a plant with increased
CC tocopherol, vitamin K, chlorophyll and/or carotenoid content. Transgenic
CC plants containing DOXS DNA coding sequences can be used for production of
CC plant and bacterial DOXS which also have increased tocopherol, vitamin K,
CC chlorophyll and/or carotenoid content. The test system can be used to
CC identify inhibitors of DOXS. This sequence represents the Streptomyces
CC avermitilis HPPD protein described in the method of the invention.
XX
SQ Sequence 306 AA:

Query Match 62.5%; Score 40; DB 21; Length 306;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPRPYLP 10
|| |||||
Db 204 rrtarphlp 213

RESULT 36
ID AAG10773
AAG10773 standard; Protein; 309 AA.
AC AAG10773;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9228.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0135621.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.5%; Score 40; DB 21; Length 309;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPPLPR 11
Db 155 rnprrpcplpk 174

RESULT 37
AAG42803
ID AAG42803 standard; Protein; 309 AA.
XX
AC AAG42803;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53423.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142053.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 62.5%; Score 40; DB 21; Length 324;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 2 RRRPPYLPK 11
Db 180 mrrppclpk 189
RESULT 39
AAC42802
ID AAC42802 standard; Protein; 324 AA.
AC AAC42802;
XX 18-OCT-2000 (first entry)
DT XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53422.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX EPI033405-A2.
FN
PD
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 23-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0144884.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		

Query Match

Best local Similarity

Matches 6; Conservative

62.5%; Score 40; DB 21; Length 324;

Pred. No. 1.9e+02;

2; Mismatches

2; Indels

0; Gaps

0;

QY 2 RRRPPYLP 11
| | | | : | :
Db 180 rnrppclpk 189

RESULT 40
ID AAG42801
AAG42801 standard; Protein: 333 AA.
XX AAG42801;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 53421.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 23-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 62.5%; Score 40; DB 21; Length 333;
Best Local Similarity 60.0%; Pred. No. 19e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 2 RRRPPYLR 11
 11111-1:
Db 189 rnrppcplpk 198

RESULT 41
AAG10771

ID	AAG10771 standard; Protein; 334 AA.
XX	
AC	AAG10771;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 9226.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 200DEP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130049.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0133284.
PR	05-MAY-1999; 99US-0133285.
PR	06-MAY-1999; 99US-0133286.
PR	07-MAY-1999; 99US-0133287.
PR	07-MAY-1999; 99US-0133288.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134421.
PR	14-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	17-JUN-1999; 99US-0139453.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.5%; Score 40; DB 21; Length 334;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRPRPYLPR 11
I I I I I : I I
Db 190 rnrprcpk 199

RESULT 42
AAR74172
ID AAR74172 standard; Protein: 421 AA.
XX AAR74172;
AC
XX 11-NOV-1995 (first entry)
DT
XX
DE Tyrosinase and tyrosinase activator protein fusion enzyme.

```

XX Plasmid pBGC646: tyrosinase; tyrosinase activator protein;
KW fusion enzyme; Streptomyces antibioticus; ORF438; melanin;
RW Escherichia coli; cloning; vector.
XX
XX Streptomyces antibioticus.
XX
XX Key Location/Qualifiers
XX Misc-difference 422 /note= "residue not indicated in coding sequence"
XX
XX MO9513386-A.
XX
XX 18-MAY-1995.
XX
XX 08-NOV-1994; 94WO-US12857.
XX
XX 12-NOV-1993; 93US-0152483.
XX
XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX
XX Della-cioppa G, Kumagai MH;
XX
XX MPI: 1995-194104/25.
XX
XX N-PSDB; AAQ92325.
XX
XX DNA encoding fusion enzyme including tyrosinase and tyrosinase
XX activator protein - useful for in vivo or in vitro prodn. of
XX melanin(s)
XX
XX Example 2; Page 30-31; 54pp; English.
XX
XX The sequence represents a tyrosinase and tyrosinase activator
XX protein fusion enzyme, encoded by an insert in plasmid pBGC646.
XX The Streptomyces antibioticus tyrosinase gene from plasmid pBGC188NDc
XX is modified and inserted in plasmid pBGC623, containing ORF438
XX (encoding a tyrosinase activator protein), to form pBGC646.
XX Transformation of Escherichia coli with the vector gives a melanin-
XX positive phenotype. The fusion enzyme induces or increases melanin
XX production in melanin-negative strains, and may also be used for in
XX vitro melanin production.
XX
XX Sequence 421 AA:
SQ

```

```

Query Match 62.5%; Score 40; DB 16; Length 421;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 RRRPRPYLP 10
DB 228 rthpspylp 237

```

RESULT 43

```

AAR74173
ID AAR74173 standard; Protein; 426 AA.
XX
XX AAR74173;
XX
XX 11-NOV-1995 (first entry)
XX
XX Tyrosinase and tyrosinase activator protein fusion enzyme.
XX
XX Plasmid pBGC648; tyrosinase; tyrosinase activator protein;
XX fusion enzyme; Streptomyces antibioticus; ORF438; melanin;
XX Escherichia coli; cloning; vector.
XX
XX Streptomyces antibioticus.
XX
XX Key Location/Qualifiers
XX Misc-difference 426 /note= "residue not indicated in coding sequence"
XX

```

```

XX
XX PN MO9513386-A.
XX
XX PD 18-MAY-1995.
XX
XX 08-NOV-1994; 94WO-US12857.
XX
XX 12-NOV-1993; 93US-0152483.
XX
XX (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX
XX Della-cioppa G, Kumagai MH;
XX
XX MPI: 1995-194104/25.
XX
XX N-PSDB; AAQ92326.
XX
XX DNA encoding fusion enzyme including tyrosinase and tyrosinase
XX activator protein - useful for in vivo or in vitro prodn. of
XX melanin(s)
XX
XX Example 3; Page 35-36; 54pp; English.
XX
XX The sequence represents a Streptomyces antibioticus tyrosinase and
XX tyrosinase activator protein fusion enzyme encoded by an insert in
XX plasmid pBGC648. The fusion enzyme has a mol.wt. 46,230, with a
XX single His residue between the 2 enzymes. Transformation of
XX Escherichia coli with the vector gives a melanin-positive phenotype.
XX Activity of the fusion enzyme encoded by pBGC648 is higher than
XX that from plasmid pBGC646 (AAQ92325). The fusion enzyme induces or
XX increases melanin production in melanin-negative strains, and may
XX also be used for in vitro melanin production.
XX
XX Sequence 426 AA:
SQ

```

```

Query Match 62.5%; Score 40; DB 16; Length 426;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 RRRPRPYLP 10
DB 380 rthpspylp 389

```

RESULT 44

```

AAR74174
ID AAR74174 standard; Protein; 478 AA.
XX
XX AAR74174;
XX
XX 11-NOV-1995 (first entry)
XX
XX Chloroplast transit peptide, tyrosinase activator protein and
XX tyrosinase fusion protein.
XX
XX Tyrosinase; tyrosinase activator protein; fusion enzyme;
XX Streptomyces antibioticus; ORF438; chloroplast transit peptide;
XX tobaco; ribulosebiphosphate-carboxylase; transgenic plant;
XX vector.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 1..37 /note= "Chloroplast transit peptide"
XX Misc-difference 478 /note= "Residue not indicated in coding sequence"
XX
XX PN MO9513386-A.
XX
XX PD 18-MAY-1995.
XX
XX 08-NOV-1994; 94WO-US12857.
XX

```

XX 12-NOV-1993; 9305-0152483.
 PR (BIOS-) BIOSOURCE TECHNOLOGIES INC.
 PA Della-cioppa G, Kumagai MH;
 PI WPI; 1995-194104/25.
 XX N-PSDB; AAQ92327.
 DR
 XX
 PT DNA encoding fusion enzyme including tyrosinase and tyrosinase
 PT activator protein - useful for in vivo or in vitro prodn. of
 PT melanin(s)
 PS
 XX Example 6; Page 37-38; 54pp; English.
 CC The sequence represents a fusion protein of a tyrosinase,
 CC tyrosinase activator protein and a chloroplast transit peptide
 CC (CTP) from tobacco ribulosebiphosphate-carboxylase (mol.wt
 CC 51,461). The CTP is fused to the N-terminus of the tyrosinase
 CC activator protein. The resulting fusion protein may be
 CC imported into higher plant chloroplasts, followed by transit
 CC peptide cleavage to give an active enzyme. Expression gives a
 CC melanin-positive phenotype.
 CC
 SQ Sequence 478 AA;

Query Match 62.5%; Score 40; DB 16; Length 478;
 Best Local Similarity 70.0%; Pred. No. 2.7e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 RRRRPPYLP 10
 ||| |||
 DB 432 rthpsspylp 441

RESULT 45
 AAB20574
 ID AAB20574 standard; Protein; 506 AA.
 XX
 AC AAB20574;
 DT 08-DEC-2000 (first entry)
 XX
 DE Perilla frutescens flavone synthase SEQ ID NO:8.
 XX
 KW Flavone; flavanone; flavone synthase; plant; flower colour;
 KW bacterial resistance; nitrogen fixing bacteria; protection;
 KW ultraviolet radiation.
 XX
 OS Perilla frutescens.
 OS
 PN WO200044907-A1.
 PN
 PD 03-AUG-2000.
 PD
 PE 28-JAN-2000; 2000WO-JP00490.
 PE
 PR 29-JAN-1999; 99JP-0022427.
 PR 19-JUL-1999; 99JP-0205229.
 XX
 PA (SUNR) SUNTORI LTD.
 PA
 PI Mizutani M, Tanaka Y, Kusumi T, Ayabe S, Akashi T;
 XX
 DR WPI; 2000-543394/49.
 DR N-PSDB; AAA88000.
 XX
 PT Gene encoding a flavone synthase for production of Anthrathum and
 PT other plants with modified flower colour
 XX
 PS Claim 3; Page 45-48; 54pp; Japanese.

XX
 CC The present sequence represents a protein which has flavone synthase
 CC activity for converting flavanone directly into flavone. Flavone
 CC synthase can be used in the production of plant varieties with modified
 CC flower colour, improved resistance to bacteria, improved interaction
 CC with nitrogen fixing bacteria, or increased protection against
 CC ultraviolet radiation.
 CC
 SQ Sequence 506 AA;

Query Match 62.5%; Score 40; DB 21; Length 506;
 Best Local Similarity 70.0%; Pred. No. 2.8e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RRRRPPYLP 10
 ||| ||| |
 DB 21 rkrpypyp 30

Search completed: September 24, 2001, 10:08:36
 Job time: 196 sec

Mon Sep 24 10:12:24 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:06:40 : Search time 26.42 Seconds
(without alignments)
8.573 Million cell updates/secTitle: US-09-276-868-4
Perfect score: 64
Sequence: 1 RRRPPPLPR 11Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues 197339

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summariesDatabase : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilestl.pep:*pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	64	100.0	14	US-09-024-975-4	Sequence 4, Appli
2	64	100.0	26	US-08-419-066-2	Sequence 2, Appli
3	64	100.0	26	US-09-024-975-2	Sequence 2, Appli
4	64	100.0	39	US-08-162-052-1	Sequence 1, Appli
5	64	100.0	39	US-08-310-722-1	Sequence 1, Appli
6	64	100.0	39	US-08-419-066-1	Sequence 1, Appli
7	64	100.0	39	US-08-728-333-1	Sequence 1, Appli
8	64	100.0	39	US-09-024-975-1	Sequence 1, Appli
9	64	100.0	39	PCT-US95-12080-1	Sequence 9, Appli
10	47	73.4	20	US-09-024-975-9	Sequence 3, Appli
11	47	73.4	59	PCT-US95-12080-2	Sequence 2, Appli
12	44	68.8	575	US-08-913-805A-10	Sequence 10, Appli
13	44	68.8	575	US-08-913-805A-10	Sequence 3, Appli
14	40	62.5	422	US-08-152-483B-3	Sequence 7, Appli
15	40	62.5	426	US-08-152-483B-7	Sequence 9, Appli
16	40	62.5	478	US-08-152-483B-9	Sequence 5, Appli
17	39	60.9	180	US-09-187-331-5	Sequence 1, Appli
18	39	60.9	195	US-09-187-331-1	Sequence 4, Appli
19	39	60.9	881	US-09-413-814-44	Sequence 5, Appli
20	38	59.4	95	US-08-911-319A-5	Sequence 5, Appli
21	38	59.4	95	US-09-352-619-5	Sequence 5, Appli
22	38	59.4	520	US-09-091-432-2	Sequence 2, Appli
23	37.5	58.6	454	US-08-764-870-5	Sequence 5, Appli
24	37.5	58.6	454	5260432-2	Patent No. 5260432
25	37	57.8	15	US-07-978-895-9	Sequence 9, Appli
26	37	57.8	15	US-08-473-119-9	Sequence 9, Appli
27	37	57.8	15	US-08-475-352-9	Sequence 9, Appli

28	37	57.8	17	US-08-630-916A-95	Sequence 95, Appli
29	37	57.8	26	US-09-024-975-8	Sequence 8, Appli
30	37	57.8	297	US-08-580-545B-6	Sequence 6, Appli
31	37	57.8	297	US-07-262-653A-6	Sequence 6, Appli
32	37	57.8	516	US-08-379-580-4	Sequence 7, Appli
33	37	57.8	581	US-08-989-386-7	Sequence 4, Appli
34	37	57.8	904	US-08-632-537-1	Sequence 1, Appli
35	37	57.8	904	PCT-US96-05316-1	Sequence 1, Appli
36	37	57.8	1047	US-08-190-687B-8	Sequence 8, Appli
37	37	57.8	1342	US-07-978-895-4	Sequence 8, Appli
38	37	57.8	1342	US-08-484-438-9	Sequence 4, Appli
39	37	57.8	1342	US-08-473-119-4	Sequence 4, Appli
40	37	57.8	1342	US-08-475-352-4	Sequence 4, Appli
41	37	57.8	1343	US-08-227-536-2	Patent No. 5183884
42	37	57.8	2414	PCT-US95-04682-2	Sequence 2, Appli
43	37	57.8	2414	US-08-602-999A-389	Sequence 389, App
44	36	56.2	16	US-08-256-747C-45	Sequence 45, Appli
45	36	56.2	29	US-08-256-747C-45	Sequence 45, Appli

ALIGNMENTS

RESULT 1
US-09-024-975-4 : Application US/09024975
Sequence 4, Appli
Patent No. 6133233
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
INVENTOR: BIERCH, FRANK
ATTORNEY: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-4

Query Match 100.0%; Score 64; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0024; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 RRRPPPLPR 11

Db

1 RRRRPPYLP 11

RESULT 2

US-08-419-066-2
Sequence 2, Application US/08419066
Patent No. 5830393

GENERAL INFORMATION:

APPLICANT: Blecha, Frank

TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDE

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: John M. Collins, Hovey, Williams, Timmons &
STREET: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/419,066

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Collins, John M.

REGISTRATION NUMBER: 26262

TELECOMMUNICATION INFORMATION:

TELEPHONE: (816) 474-9050

TELEFAX: (816) 474-9057

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 26 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

US-08-419-066-2

Query Match

Best Local Similarity 100.0%; Score 64; DB 2; Length 26;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPYLP 11

Db 1 RRRRPPYLP 11

RESULT 3

STATE: MO
COUNTRY: USA
ZIP: 64108

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/024,975

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/802,306

FILING DATE: 18-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: COLLINS, JOHN M.

REGISTRATION NUMBER: 26,262

TELECOMMUNICATION INFORMATION:

TELEPHONE: 816/474-9050

TELEFAX: 816/474-9057

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 26 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-024-975-2

Query Match

Best Local Similarity 100.0%; Score 64; DB 4; Length 26;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPYLP 11

Db 1 RRRRPPYLP 11

RESULT 4

US-08-162-052-1

Sequence 1, Application US/08162052

Patent No. 5489575

GENERAL INFORMATION:

APPLICANT: LEE, Jong-Youn

APPLICANT: BOMAN, Hans G

APPLICANT: MTT, Viktor

TITLE OF INVENTION: NOVEL POLYPEPTIDES AND THEIR USE

ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 003300-299
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-162-052-1

Query Match 100.0%; Score 64; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPYLP 11
DB 1 RRRRPPYLP 11

RESULT 5
US-08-310-722-1
Sequence 1, Application US/08310722
Patent No. 5654273
GENERAL INFORMATION:
APPLICANT: Gallo, Richard L.
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/310,722
FILING DATE: 22-SEP-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: CMCC379
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Mutt, Viktor
AUTHORS: Jorhval, Hans
TITLE: No. 5654273el Polypeptides And Their Use
JOURNAL: PCT WO 92/22578

DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
US-08-310-722-1

Query Match 100.0%; Score 64; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPYLP 11
DB 1 RRRRPPYLP 11

RESULT 6
US-08-419-066-1
Sequence 1, Application US/08419066
Patent No. 5830993
GENERAL INFORMATION:
APPLICANT: Blecha, Frank
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: John M. Collins, Hovey, Williams, Timmons &
STREET: 2405 Grand Boulevard, Suite 400
CITY: Kansas City
STATE: Missouri
COUNTRY: U.S.A.
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,066
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 23625
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816) 474-9050
TELEFAX: (816) 474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-419-066-1

Query Match 100.0%; Score 64; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPYLP 11
DB 1 RRRRPPYLP 11

RESULT 7
US-08-728-333-1
Sequence 1, Application US/08728333
Patent No. 5863897

GENERAL INFORMATION:
APPLICANT: Gallo, Richard L.
APPLICANT: Klagsbrun, Michael
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,333
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/310,722
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: CMCC379
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Mutt, Viktor
AUTHORS: Jornvall, Hans
TITLE: No. 5863897el Polypeptides And Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
US-08-728-333-1

Query Match 100.0%; Score 64; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRPRPYLPR 11
Db 1 RRRPRPYLPR 11

RESULT 8
US-09-024-975-1
Sequence 1, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400

CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-1

Query Match 100.0%; Score 64; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRPRPYLPR 11
Db 1 RRRPRPYLPR 11

RESULT 9
PCT-US95-12080-1
Sequence 1, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Mutt, Viktor
AUTHORS: Jorvall, Hans
TITLE: Novel Polypeptides And Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
PCT-US95-12080-1

Query Match 100.0%; Score 64; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPPLPR 11
Db 1 RRRPPPLPR 11

RESULT 10
US-09-024-975-9
Sequence 9, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-9

Query Match 73.4%; Score 47; DB 4; Length 20;
Best Local Similarity 81.8%; Pred. No. 0.7;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPPLPR 11
Db 2 RRRPPPLPR 12

RESULT 11
PCT-US95-12080-3
Sequence 3, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-3

Query Match 73.4%; Score 47; DB 5; Length 59;
Best Local Similarity 81.8%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPPLPR 11
Db 2 RRRPPPLPR 12

RESULT 12
US-08-913-805A-2
Sequence 2, Application US/08913805A
Patent No. 6054304
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: UOZUMI, Naofumi
APPLICANT: SHIBA, Tetsuo
APPLICANT: YAMAGIDANI, Shusaku
TITLE OF INVENTION: Alpha 1-6 Fucosyltransferase
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3+ Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: WordPerfect 6.1 Windows
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,805A
FILING DATE: 7 JAN 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/00171
FILING DATE: 23 JAN 1997
APPLICATION NUMBER: JP 192260
FILING DATE: 22 JUL 1996
APPLICATION NUMBER: JP 162813
FILING DATE: 24 JUN 1996
APPLICATION NUMBER: JP 161648
FILING DATE: 21 JUN 1996
APPLICATION NUMBER: JP 10365
FILING DATE: 24 JAN 1996
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2356/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-805A-2

Query Match 68.8%; Score 44; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PRPYLP 10
DB 299 PRPYLP 305

RESULT 13
US-08-913-805A-10
Sequence 10, Application US/08913805A
Patent No. 6054304
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: UOZUMI, Naofumi
APPLICANT: SHIBA, Tetsuo
APPLICANT: YAMAGIDANI, Shusaku
TITLE OF INVENTION: Alpha 1-6 Fucosyltransferase
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3+ Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: Wordperfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,805A
FILING DATE: 7 JAN 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/00171
FILING DATE: 23 JAN 1997
APPLICATION NUMBER: JP 192260
FILING DATE: 22 JUL 1996
APPLICATION NUMBER: JP 162813
FILING DATE: 24 JUN 1996

APPLICATION NUMBER: JP 161648
FILING DATE: 21 JUN 1996
APPLICATION NUMBER: JP 10365
FILING DATE: 24 JAN 1996
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2356/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-913-805A-10

Query Match 68.8%; Score 44; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PRPYLP 10
DB 299 PRPYLP 305

RESULT 14
US-08-152-483B-3
Sequence 3, Application US/08152483B
Patent No. 5529909
GENERAL INFORMATION:
APPLICANT: Della-Cioppa, Guy
APPLICANT: Kumagai, Kento
TITLE OF INVENTION: TYROSINASE-ACTIVATOR
TITLE OF INVENTION: PROTEIN FUSION ENZYME
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 2730 Sand Hill Road
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,483B
FILING DATE: No. 5529909ember 12, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7/857,602
FILING DATE: March 30, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 923,692
FILING DATE: July 31, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 600,244
FILING DATE: October 22, 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 641,617
FILING DATE: January 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 737,899
FILING DATE: July 26, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.

REGISTRATION NUMBER: 25,227
 REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 854-3660
 TELEFAX: (415) 854-3694
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 422
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: NO
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM:
 IMMEDIATE SOURCE:
 CLONE:
 FEATURE:
 US-08-152-483B-3

Query Match 62.5%; Score 40; DB 1; Length 422;
 Best Local Similarity 70.0%; Pred. No. 1.le+02; 3; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 RRRPPPLP 10
 ||| |||
 Db 228 RHPSPPLP 237

RESULT 15
 US-08-152-483B-7
 ; Sequence 7, Application US/08152483B
 ; Patent No. 5529909
 ; GENERAL INFORMATION:
 ; APPLICANT: Della-Cloppa, Guy
 ; APPLICANT: Kumagai, Monto
 ; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
 ; TITLE OF INVENTION: PROTEIN FUSION ENZYME
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 2730 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0,
 ; SOFTWARE: Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/152,483B
 ; FILING DATE: No. 5529909 September 12, 1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 7/857,602
 ; FILING DATE: March 30, 1992
 ; PRIOR APPLICATION DATA: 923,692
 ; APPLICATION NUMBER: 923,692
 ; FILING DATE: July 31, 1992
 ; PRIOR APPLICATION DATA: 600,244
 ; APPLICATION NUMBER: 600,244
 ; FILING DATE: October 22, 1990
 ; PRIOR APPLICATION DATA: 641,617
 ; APPLICATION NUMBER: 641,617
 ; FILING DATE: January 16, 1991
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 737,899
 FILING DATE: July 26, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Halluin, Albert P.
 REGISTRATION NUMBER: 25,227
 REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 854-3660
 TELEFAX: (415) 854-3694
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 426
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: circular
 MOLECULE TYPE: protein
 DESCRIPTION: NO
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM:
 IMMEDIATE SOURCE:
 CLONE:
 FEATURE:
 US-08-152-483B-7

Query Match 62.5%; Score 40; DB 1; Length 426;
 Best Local Similarity 70.0%; Pred. No. 1.le+02; 3; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 RRRPPPLP 10
 ||| |||
 Db 380 RHPSPPLP 389

RESULT 16
 US-08-152-483B-9
 ; Sequence 9, Application US/08152483B
 ; Patent No. 5529909
 ; GENERAL INFORMATION:
 ; APPLICANT: Della-Cloppa, Guy
 ; APPLICANT: Kumagai, Monto
 ; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
 ; TITLE OF INVENTION: PROTEIN FUSION ENZYME
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 2730 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0,
 ; SOFTWARE: Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/152,483B
 ; FILING DATE: No. 5529909 September 12, 1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 7/857,602
 ; FILING DATE: March 30, 1992
 ; PRIOR APPLICATION DATA: 923,692
 ; APPLICATION NUMBER: 923,692
 ; FILING DATE: July 31, 1992
 ; PRIOR APPLICATION DATA: 600,244
 ; APPLICATION NUMBER: 600,244
 ; FILING DATE: October 22, 1990
 ; PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 641,617
FILING DATE: January 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 737,899
FILING DATE: July 26, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Halliuh, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-3660
TELEFAX: (415) 854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 478
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM:
IMMEDIATE SOURCE:
CLONE:
FEATURE:
US-08-152-483B-9

Query Match 62.5%; Score 40; DB 1; Length 478;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPYLP 10
DB 432 RRRHSSPYLP 441

RESULT 17
US-09-187-331-5
Sequence 5, Application US/09187331
Patent No. 6043056
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
FILE REFERENCE: PF-0631 US
CURRENT APPLICATION NUMBER: US/09/187,331
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PERL Program
SEQ ID NO 5
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 92499136
US-09-187-331-5

Query Match 60.9%; Score 39; DB 3; Length 180;
Best Local Similarity 55.6%; Pred. No. 67;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPPYLP 11
DB 47 KKPPYYPQ 55

RESULT 18
US-09-187-331-1
Sequence 1, Application US/09187331
Patent No. 6043056
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
FILE REFERENCE: PF-0631 US
CURRENT APPLICATION NUMBER: US/09/187,331
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 195
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 2297891
US-09-187-331-1

Query Match 60.9%; Score 39; DB 3; Length 195;
Best Local Similarity 55.6%; Pred. No. 72;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPPYLP 11
DB 47 KKPPYYPQ 55

RESULT 19
US-09-413-814-44
Sequence 44, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M.
APPLICANT: Dougherty, Brian A.
APPLICANT: Goldberg, Steven L.
APPLICANT: Hoile, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 881
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-44

Query Match 60.9%; Score 39; DB 4; Length 881;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPYLP 10

Db 597 KGRROPVLP 606

RESULT 20

US-08-911-319A-5
Sequence 5, Application US/08911319A
Patent No. 5968798
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,319A
FILING DATE: August 14, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Muenzen, Colette C.
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0363 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: A-Gene Seq
CLONE: PJ002 PAR19P
US-08-911-319A-5

Query Match 59.4%; Score 38; DB 2; Length 95;
Best Local Similarity 62.5%; Pred. No. 50;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 4 PRPPVLP 11
Db 75 PKPPLPK 82

RESULT 21

US-09-352-619-5
Sequence 5, Application US/09352619
Patent No. 6084070
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/352,619
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Muenzen, Colette C.
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0363 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: A-Gene Seq
CLONE: PJ002 PAR19P
US-09-352-619-5

Query Match 59.4%; Score 38; DB 3; Length 95;
Best Local Similarity 62.5%; Pred. No. 50;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 PRPPVLP 11
Db 75 PKPPLPK 82

RESULT 22
US-09-091-432-2
Sequence 2, Application US/09091432
Patent No. 5961837
GENERAL INFORMATION:
APPLICANT: Chapple, Clint
TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
FILE REFERENCE: 7024-325
CURRENT APPLICATION NUMBER: US/09/091,432
CURRENT FILING DATE: 1998-06-18
EARLIER APPLICATION NUMBER: PCT/US96/20094
EARLIER FILING DATE: 1996-12-19
EARLIER APPLICATION NUMBER: US 60/009,119
EARLIER FILING DATE: 1995-12-22
EARLIER APPLICATION NUMBER: US 60/013,388
EARLIER FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Microsoft Word 2.0C
SEQ ID NO 2
LENGTH: 520
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Sequence is deduced from DNA sequence of SEQ ID NO:1
Patent No. 5961837
US-09-091-432-2

Query Match 59.4%; Score 38; DB 2; Length 520;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 RRRPPYLP 10
Db 34 RRRPPYLP 42

RESULT 23
US-08-764-870-5
Sequence 5, Application US/08764870
Patent No. 6236946
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S
APPLICANT: Baxter, John D
APPLICANT: Fletcher, Robert J
APPLICANT: Wagner, Richard L
APPLICANT: Kushner, Peter J
APPLICANT: Applell, James W
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
TITLE OF INVENTION: Binding Domains
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward
STREET: Five Palo Alto Square, 3000 El Camino Real
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,870
FILING DATE: 13-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,543
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N
REGISTRATION NUMBER: 35,966
REFERENCE/DOCKET NUMBER: UCAL-246/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)843-5000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 454 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-764-870-5

Query Match 58.6%; Score 37.5; DB 4; Length 454;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 RRRPPYLP-PR 11
Db 367 RRRPPYLP-PR 378

RESULT 24

5260432-2
Patent No. 5260432
APPLICANT: TAKAKU, FUMIMARO; ISHIKAWA, TAKASHI; IWAMARI, MICHIO;
EVANS, RONALD M.; UMESONO, KAZUHIKO
TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/370,407
FILING DATE: 22-JUN-1989
SEQ ID NO: 2
LENGTH: 454
5260432-2

Query Match 58.6%; Score 37.5; DB 6; Length 454;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 1 RRRPPYLP-PR 11
Db 367 RRRPPYLP-PR 378

RESULT 25
US-07-978-895-9
Sequence 9, Application US/07978895
Patent No. 5480968
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/978,895
FILING DATE: 19921110
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,406
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-07-978-895-9

Query Match 57.8%; Score 37; DB 1; Length 15;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRRPPYLPYR 11
 111 11:11
 Db 5 RRRHSPHPYR 15

RESULT 26

US-08-473-119-9
 ; Sequence 9, Application US/08473119
 ; Patent No. 5820859

GENERAL INFORMATION:

APPLICANT: Kraus, Mathias H.
 APPLICANT: Aaronson, Stuart A.
 TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
 TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THERETO, AND
 TITLE OF INVENTION: BIOSAYS AND METHODS RELATED THERETO
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Suite 400
 STREET: 133 Carnegie Way, N.W.
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: U.S.A.
 ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473.119
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/978, 895
 FILING DATE: 10-NOV-1992
 APPLICATION NUMBER: US 07/444,406
 FILING DATE: 01-DEC-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Perryman, David G.
 REGISTRATION NUMBER: 33,438
 REFERENCE/DOCKET NUMBER: 1414-028
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 688-0770
 TELEFAX: (404) 688-9880

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 US-08-473-119-9

Query Match 57.8%; Score 37; DB 2; Length 15;

Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRRPPYLPYR 11
 111 11:11
 Db 5 RRRHSPHPYR 15

RESULT 27

US-08-475-352-9
 ; Sequence 9, Application US/08475352
 ; Patent No. 5916755

GENERAL INFORMATION:

APPLICANT: Kraus, Mathias H.
 APPLICANT: Aaronson, Stuart A.
 TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
 TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THERETO, AND

TITLE OF INVENTION: BIOSAYS AND METHODS RELATED THERETO
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Suite 400
 STREET: 133 Carnegie Way, N.W.
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: U.S.A.
 ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/475,352
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/978, 895

FILING DATE: 07/444,406

FILING DATE: 01-DEC-1989

ATTORNEY/AGENT INFORMATION:

NAME: Perryman, David G.

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414-028

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 688-0770

TELEFAX: (404) 688-9880

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-475-352-9

Query Match 57.8%; Score 37; DB 2; Length 15;

Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRRPPYLPYR 11
 111 11:11
 Db 5 RRRHSPHPYR 15

RESULT 28

US-08-630-916A-95
 ; Sequence 95, Application US/08630916A
 ; Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio
 APPLICANT: Kay, Brian K.
 APPLICANT: Rowles, Dana M.
 TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
 NUMBER OF SEQUENCES: 124
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-916A-95

Query Match 57.8%; Score 37; DB 3; Length 17;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPPPYLP 10
1 11111
DB 8 PPPYLP 14

RESULT 29
US-09-024-975-8
Sequence 8, Application US/09024975
Patent No. 613323
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECH, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9057
TELEFAX: 816/474-9050
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-8

Query Match 57.8%; Score 37; DB 4; Length 26;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RPPPYLP 11
1111111
DB 14 RPPPYLP 22

RESULT 30
US-08-580-545B-6
Sequence 6, Application US/08580545B
Patent No. 5932713
GENERAL INFORMATION:
APPLICANT: Yoshikawa, Kasukabe
APPLICANT: Koichi, Fujisawa
APPLICANT: Susumu, Nishiguchi
APPLICANT: Yoshiko, Maekawa
APPLICANT: Randy, Allen
TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 601 Thirteenth Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,545B
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bretschneider, Barry E.
REGISTRATION NUMBER: 28,055
REFERENCE/DOCKET NUMBER: 04473/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/783-5070
TELEFAX: 202/783-2331
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-545B-6

Query Match 57.8%; Score 37; DB 2; Length 297;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RPPPYLP 10
11111
DB 172 KAPPPYMP 180

RESULT 31
US-09-262-653A-6
Sequence 6, Application US/09262653A
Patent No. 6166294
GENERAL INFORMATION:
APPLICANT: Yoshikawa, Kasukabe
APPLICANT: Koichi, Fujisawa
APPLICANT: Susumu, Nishiguchi
APPLICANT: Yoshiko, Maekawa
APPLICANT: Randy, Allen

TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 601 Thirteenth Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/262,653A
APPLICATION NUMBER: US/09/262,653A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bretschneider, Barry E.
REGISTRATION NUMBER: 28,055
REFERENCE/DOCKET NUMBER: 04473/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/783-2331
TELEFAX: 202/783-5070
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 297 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-262-653A-6

Query Match 57.8%; Score 37; DB 4; Length 297;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 2 RRRPPVLP 10
DB 172 KAPPPVMP 180

RESULT 32
US-08-379-580-4
Sequence 4, Application US/08379580
Patent No. 6180362
GENERAL INFORMATION:
APPLICANT: Duchesne, Marc
APPLICANT: Schweighoffer, Fabien
APPLICANT: Tocque, Bruno
TITLE OF INVENTION: Peptides Inhibiting Ras Protein
TITLE OF INVENTION: Activity, Preparation and Use Thereof
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/379,580
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00772
FILING DATE: 28-JUL-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/09433
FILING DATE: 03-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92049-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-379-580-4

Query Match 57.8%; Score 37; DB 4; Length 516;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 4 PPPVLP 10
DB 8 PPPVLP 14

RESULT 33
US-08-989-386-7
Sequence 7, Application US/08989386
Patent No. 5989860
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN ISOMERASE HOMOLOGS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,386
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0443 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 581 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank
CLONE: 894162
US-08-989-386-7

Query Match

Best Local Similarity 57.8%; Score 37; DB 2; Length 581;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 5 RRPYPR 11
Db 162 RPYCPR 168

RESULT 34

US-08-632-537-1
Sequence 1, Application US/08632537
Patent No. 6197497
GENERAL INFORMATION:
APPLICANT: Goade, Diane E.
APPLICANT: Bell, Richard
APPLICANT: Jenison, Steven
TITLE OF INVENTION: Immunoassay for Herpes Simplex Virus.
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Hoffman, Wasson & Gitler
STREET: 2361 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 500 Kb storage
COMPUTER: AOPEN Pentium
OPERATING SYSTEM: Wordperfect 5.1
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/632,537
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426,604
FILING DATE: 21-APR-95
ATTORNEY/AGENT INFORMATION:
NAME: Buttm, Jean A.
REGISTRATION NUMBER: 24,236
REFERENCE/DOCKET NUMBER: A5144CIP.SEO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)415-0100
TELEFAX: (703)418-2768
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 amino acids
TYPE: amino acid
TOPOLOGY: linear
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 14-110; 295-507; 814-901
US-08-632-537-1

Query Match

Best Local Similarity 57.8%; Score 37; DB 4; Length 904;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 2 RRPYPR 11
Db 73 RRPYPR 82

RESULT 35

PCT-US96-05316-1
Sequence 1, Application PC/TUS9605316

GENERAL INFORMATION:

APPLICANT: Goade, Diane E.
APPLICANT: Bell, Richard
APPLICANT: Jenison, Steven
TITLE OF INVENTION: Immunoassay for Herpes Simplex Virus.
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Hoffman, Wasson & Gitler
STREET: 2361 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 500 Kb storage
COMPUTER: AOPEN Pentium
OPERATING SYSTEM: Wordperfect 5.1
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05316
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/426,604
FILING DATE: 21-APR-95
ATTORNEY/AGENT INFORMATION:
NAME: Gitler, Stewart L.
REGISTRATION NUMBER: 31,256
REFERENCE/DOCKET NUMBER: A5144PCT.SEO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)415-0100
TELEFAX: (703)418-2768
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 amino acids
TYPE: amino acid
TOPOLOGY: linear
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 14 TO 110
RELEVANT RESIDUES IN SEQ ID NO: FROM 295 TO 507
RELEVANT RESIDUES IN SEQ ID NO: FROM 814 TO 901
PCT-US96-05316-1

Query Match

Best Local Similarity 57.8%; Score 37; DB 5; Length 904;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 2 RRPYPR 11
Db 73 RRPYPR 82

RESULT 36

US-08-190-687B-8
Sequence 8, Application US/08190687B
Patent No. 5760203
GENERAL INFORMATION:

APPLICANT: Wong, Gail L.
APPLICANT: Martin, George
APPLICANT: McCormick, Francis P.
APPLICANT: Rubinfeld, Bonnie
APPLICANT: O'Rourke, Edward C.
APPLICANT: Clark, Robin
TITLE OF INVENTION: GAP Gene Sequences
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/190,687B
APPLICATION NUMBER: 02-FEB-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/774,644
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/260,807
FILING DATE: 21-OCT-1988
PRIOR APPLICATION DATA: 07/230,761
APPLICATION NUMBER: 10-AUG-1988
FILING DATE: 10-AUG-1988
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 27527/31898
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1047 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-190-687B-8

Query Match 57.8%; Score 37; DB 1; Length 1047;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 PPPPYLP 10
111111
DB 138 PPPPYLP 144

RESULT 37
US-07-978-895-4
Sequence 4; Application US/07978895
Patent No. 5480968
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, AND
TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THERETO
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/978,895
FILING DATE: 19921110
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,406

FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-9880
TELEFAX: (404) 688-0770
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-978-895-4

Query Match 57.8%; Score 37; DB 1; Length 1342;
Best Local Similarity 63.6%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPPPYLP 11
111111
DB 1203 RRRHSPHP 1213

RESULT 38
US-08-484-438-9
Sequence 9; Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Siegall, Clay B.
APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.
TITLE OF INVENTION: HERA HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-9

Query Match 57.8%; Score 37; DB 2; Length 1342;
Best Local Similarity 63.6%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRRPPYLPR 11
Db 1203 RRRHSPHPPR 1213

RESULT 39
US-08-473-119-4
Sequence 4, Application US/08473119
Patent No. 5820859
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREO, AND
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,119
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,895
FILING DATE: 10-NOV-1992
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
FAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-473-119-4

Query Match 57.8%; Score 37; DB 2; Length 1342;
Best Local Similarity 63.6%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRRPPYLPR 11

Db 1203 RRRHSPHPPR 1213

RESULT 40
US-08-475-352-4
Sequence 4, Application US/08475352
Patent No. 5916755
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREO, AND
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,352
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,895
FILING DATE:
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
FAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-475-352-4

Query Match 57.8%; Score 37; DB 2; Length 1342;
Best Local Similarity 63.6%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRRPPYLPR 11
Db 1203 RRRHSPHPPR 1213

RESULT 41
US-08-475-352-4
Sequence 4, Application US/08475352
Patent No. 5916755
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.; AARONSON, STUART A.
TITLE OF INVENTION: DNA SEGMENT ENCODING A GENE FOR A
RECEPTOR RELATED TO THE EPIDERMAL GROWTH FACTOR RECEPTOR
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/444,406
FILING DATE: 01-DEC-1989
SEQ ID NO: 4:
LENGTH: 1343

FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 389:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-389

Query Match 56.2%; Score 36; DB 4; Length 16;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

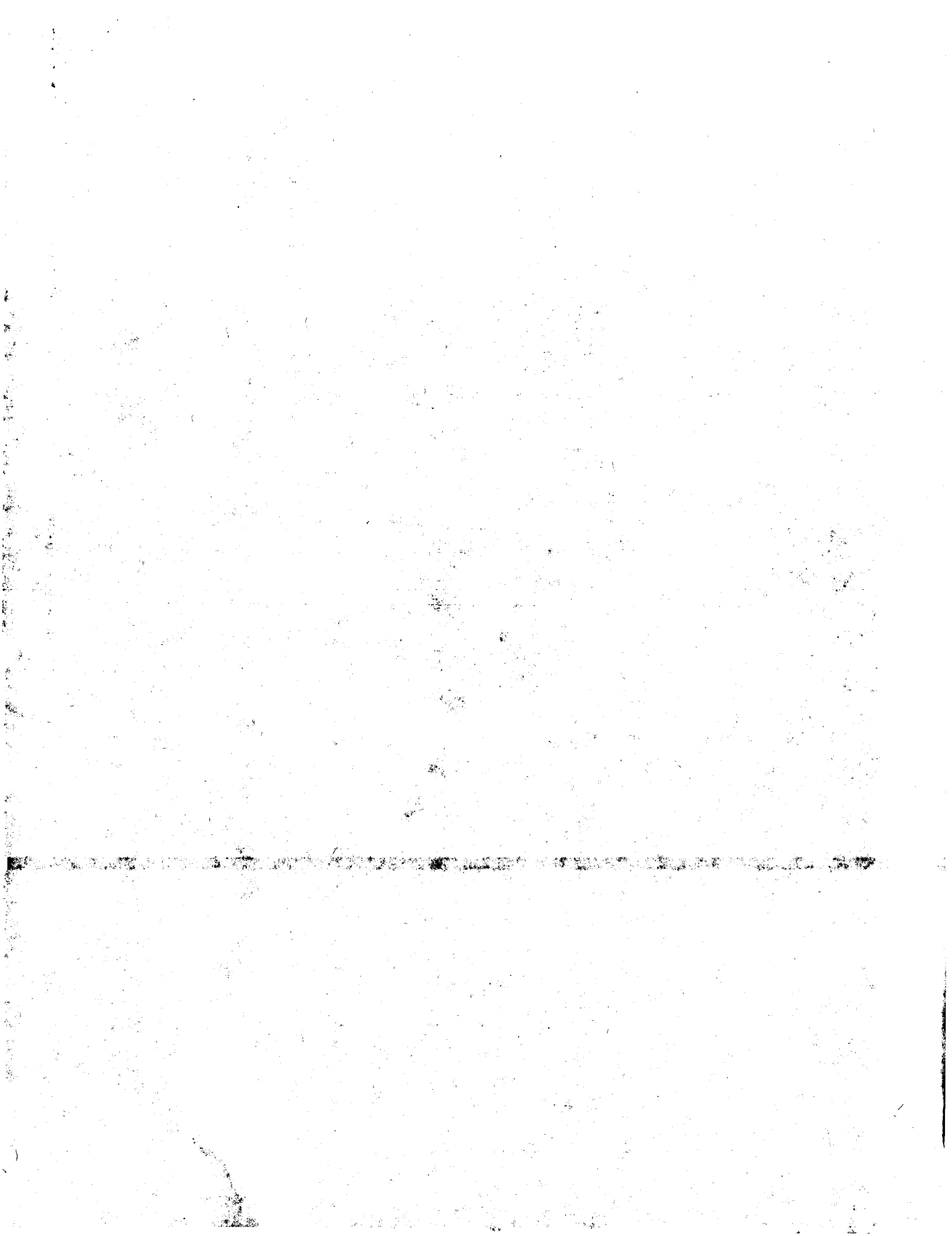
QY 2 RRRPPYP 10
:||||| 1
DB 4 QRRPPQKP 12

RESULT 45
US-08-256-747C-45
Sequence 45, Application US/08256747C
Patent No. 6037448
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: GILLAM, Shirley
APPLICANT: OU, Dawei
APPLICANT: TINGLE, Audrey
TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,747C
FILING DATE: 06-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-370
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-747C-45

Query Match 56.2%; Score 36; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 31;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRRPP 7
:||||| 1
DB 8 RRRPP 13

Search completed: September 24, 2001, 10:06:41
Job time: 81 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:06:03 ; Search time 29.48 Seconds
(without alignments)
28.423 Million cell updates/sec

Title: US-09-276-868-4
Perfect score: 64
Sequence: 1 RRRRPPPLPR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	64	100.0	172 2	S68232 antimicrobial prot
2	47	73.4	59 2	A36589 bacitracin 7 - bov
3	46	71.9	184 2	T29373 hypothetical prote
4	45	70.3	354 2	C70583 hypothetical prote
5	45	70.3	372 2	S69501 DNA-binding protei
6	44	68.8	82 2	A41051 spore coat protein
7	44	68.8	575 1	JC5432 glycoprotein 6-alp
8	42	65.6	589 2	T29299 hypothetical prote
9	42	65.6	740 2	S43566 R01H10.8 protein (
10	41	64.1	104 2	G84607 hypothetical prote
11	41	64.1	261 1	MMBXEXE infected cell prot
12	41	64.1	665 2	D96621 auxin response fac
13	40	62.5	129 2	T03861 glycine-rich prote
14	40	62.5	262 2	A70577 hypothetical prote
15	40	62.5	273 2	B23971 monophenol monooxy
16	40	62.5	301 2	JQ1663 hybrid proline-ric
17	40	62.5	324 2	T01316 epoxide hydrolase
18	40	62.5	439 2	I60916 HNF-3/forkhead hom
19	40	62.5	587 2	S63033 hypothetical prote
20	40	62.5	600 2	G83081 conserved hypochet
21	40	62.5	693 2	JN0674 ubiquitin-like fus
22	40	62.5	701 2	JN0674 ubiquitin-like fus
23	40	62.5	963 2	T19140 hypothetical prote
24	39	60.9	134 2	E72532 hypothetical prote
25	39	60.9	148 1	Q0BE20 BLRP protein - hu
26	39	60.9	180 2	S43791 PBDX protein - hum
27	39	60.9	190 2	S68230 conserved hypochet
28	39	60.9	224 2	G82780 related to DNA-dir
29	39	60.9	333 2	T49750

30	39	60.9	380 2	B82634 threonine dehydrat
31	39	60.9	467 2	G85957 unknown protein en
32	39	60.9	488 2	T33739 hypothetical prote
33	39	60.9	505 2	C96817 F9K20.15 (imported
34	39	60.9	537 2	F85605 unknown in ISEC8 l
35	39	60.9	537 2	G85656 unknown in ISEC8 l
36	39	60.9	772 2	T27512 hypothetical prote
37	38.5	60.2	454 2	S06124 retinoic acid rece
38	38	59.4	130 2	S75341 hypothetical prote
39	38	59.4	139 4	S09612 hypothetical xrepB
40	38	59.4	215 2	A03863 hypothetical prote
41	38	59.4	237 2	T43472 hypothetical prote
42	38	59.4	245 2	S28827 chlorophyll a/b-bi
43	38	59.4	272 2	T30959 hypothetical prote
44	38	59.4	322 2	T35646 probable thiamin m
45	38	59.4	373 2	A47234 homeobox protein H

ALIGNMENTS

RESULT 1
S68232
antimicrobial protein PR-39 precursor, cathelin-associated - pig
N:Alternate names: myeloid antibacterial protein PR-39
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68232; JN0899; I47138; S19563
R:Zhao, C.; Ganz, T.; Lehrer, R.I.
FEBS Lett. 376, 130-134, 1995
A:Title: Structures of genes for two cathelin-associated antimicrobial peptides: prop
A:Reference number: S68232; MUID:96105365
A:Accession: S68232
A:Molecule type: DNA
A:Status: translation not shown
A:Residues: 1-172 <ZHA>
A:Cross-references: EMBL:X89201; NID:g1165150; PIDN:CAA61487.1; PID:g1165151
A:Experimental source: leukocytes
R:Storici, P.; Zanetti, M.
Biochem. Biophys. Res. Commun. 196, 1058-1065, 1993
A:Title: A cDNA derived from pig bone marrow cells predicts a sequence identical to t
A:Reference number: JN0899; MUID:94071853
A:Accession: JN0899
A:Molecule type: mRNA
A:Residues: 1-20, 'A', 22-172 <STO>
A:Cross-references: GB:I23825; NID:g435100; PIDN:AAA31109.1; PID:g435101
A:Experimental source: bone marrow cells
R:Gundundsson, G.H.; Magnusson, K.P.; Chowdhary, B.P.; Johansson, M.; Andersson, L.;
Proc. Natl. Acad. Sci. U.S.A. 92, 7085-7089, 1995
A:Title: Structure of the gene for porcine peptide antibiotic PR-39, a cathelin gene
A:Reference number: I47138; MUID:95350216
A:Accession: I47138
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28, 'T', 30-89, 'OR', 92-116, 'NDP', 120-172 <GUD>
A:Cross-references: EMBL:X87236; NID:g829142; PIDN:CAA60682.1; PID:g1051298
R:Agarwal, B.; Lee, J.Y.; Bergman, T.; Carlquist, M.; Boman, H.G.; Mutt, V.; Joernv
Eur. J. Biochem. 202, 849-854, 1991
A:Title: Amino acid sequence of PR-39. Isolation from pig intestine of a new member o
A:Reference number: S19563; MUID:92111534
A:Accession: S19563
A:Molecule type: protein
A:Residues: 131-169 <AGE>
A:Experimental source: Intestine
C:Gene: PR39
A:Gene: PR39
A:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end; antibacterial
F:1-29/Domain: signal sequence #Status predicted <SIG>
F:22-129/Domain: cystatin homology <CYS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-169/Product: antimicrobial protein PR-39 #status experimental <MAT>

F:169/Modified site: amidated carboxyl end (Pro) (amide in mature form from following gl

Query Match 100.0%; Score 64; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
| | | | | | | | | |
DB 131 RRRPPPYLPR 141

RESULT 2

A36589

bactenein 7 - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 09-May-1997

C:Accession: A36589

R:Frank, R.W.; Gennaro, R.; Schneider, K.; Przybylski, M.; Romeo, D.

J. Biol. Chem. 265, 18871-18874, 1990

A:Title: Amino acid sequences of two proline-rich bactericins. Antimicrobial peptides of

A:Reference number: A36589; MUID:91035404

A:Accession: A36589

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-59 <FRA>

C:Superfamily: cathelin; cystatin homology

Query Match 73.4%; Score 47; DB 2; Length 59;
Best Local Similarity 81.8%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 11
| | | | | | | |
DB 2 RRRPPPYLPR 12

RESULT 3

T29373

hypothetical protein ZC404.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T29373

R:Bentley, D.; Le, T.T.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid ZC404.

A:Reference number: Z20614

A:Accession: T29373

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-184 <BEN>

A:Cross-references: EMBL:U55363; PIDN:AAA97967.1; GSPDB:GN00023; CESP:ZC404.1

A:Experimental source: strain Bristol N2; clone ZC404

C:Genetics: CESP:ZC404.1

A:Gene: CESP:ZC404.1

A:Map position: 5

A:Introns: 15/2; 50/2; 75/2; 138/2

C:Superfamily: Caenorhabditis elegans hypothetical protein ZC404.1

Query Match 71.9%; Score 46; DB 2; Length 184;
Best Local Similarity 88.9%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPPYLPR 11
| | | | | | | | | |
DB 26 RRRPPPYLPR 34

RESULT 4

C70583

hypothetical protein RV0923c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70583

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: C70583

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-354 <COL>

A:Cross-references: GB:Z95210; GB:AL123456; NID:93261757; PIDN:CAB08505.1; PID:e31517

A:Experimental source: strain H37RV

C:Genetics: RV0923c

Query Match 70.3%; Score 45; DB 2; Length 354;
Best Local Similarity 88.9%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPPYLPR 11
| | | | | | | | | |
DB 111 RRRPPPYLPR 119

RESULT 5

S69501

DNA-binding protein A variant - human

N:Alternate names: cold shock domain protein A

C:Species: Homo sapiens (man)

C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S69501

R:Coles, L.S.; Diamond, P.; Occhiodoro, F.; Vadás, M.A.; Shannon, M.F.

Nucleic Acids Res. 24, 2311-2317, 1996

A:Title: Cold shock domain proteins repress transcription from the GM-CSF promoter.

A:Reference number: S69501; MUID:96279731

A:Accession: S69501

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-372 <COL>

A:Cross-references: EMBL:X95325; NID:q1167837; PIDN:CAA64631.1; PID:q1167838

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996

C:Superfamily: y box-binding protein 1; cold shock domain homology

F:93-157/Domain: cold shock domain homology <CSD>

Query Match 70.3%; Score 45; DB 2; Length 372;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLPR 10
| | | | | | | |
DB 336 RRRPPPNAP 345

RESULT 6

A41051

spore coat protein precursor - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 03-Apr-1992 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999

C:Accession: S04835; A41051; F69606

R:Aranson, A.I.; Song, H.Y.; Bourne, N.

Mol. Microbiol. 3, 437-444, 1989

A:Title: Gene structure and precursor processing of a novel Bacillus subtilis spore c

A:Reference number: S04835; MUID:89313296

A:Accession: S04835

A:Molecule type: DNA

A:Residues: 'MNVHTPNTLSIRNVKGIKAREVILL', 2-82 <AR2>

A:Cross-references: EMBL:X13740; NID:q39864; PIDN:CAA32004.1; PID:q39865

A:Experimental source: strain JH642
A:Note: part of this sequence, including the amino end of the mature protein, was confir
B: Bourne, N.; Fitzhames, P.C.; Aronson, A.I.
J. Bacteriol. 173, 6618-6623, 1991
A:Title: Structural and germination defects of *Bacillus subtilis* spores with altered con
A:Reference number: A41051; MUID:92011439
A:Accession: A41051
A:Molecule type: Protein
A:Residues: 311-311 <BOU>
A:Experimental source: strain JH642
A:Note: The material sequenced was the larger of two isolated precursor forms; the amin
A:Note: Both the location of the transcription start site and peptide sequencing of the
R: Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berth
C: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleicher, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zimmstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:96044033
A:Accession: F69606
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-589 <JON>
A:Cross-references: GB:259110; GB:AL009126; NID:92633472; PIDN:CAB13066.1; PID:el183229;
A:Experimental source: strain 168
C:Comment: This structural protein is expressed during stage V of sporulation.
C:Genetics:
A:Gene: cotT
A:Start codon: TTG
C:Keywords: sporulation
F:1-19/Domain: propeptide #status experimental <PRO>
F:20-82/Product: spore coat protein #status experimental <MAT>
Query Match 68.8%; Score 44; DB 2; Length 82;
Best Local Similarity 87.5%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 RRRPPYLP 10
Db 57 RRRPPYLP 64
RESULT 7
JC5432
glycoprotein 6-alpha-L-fucosyltransferase (EC 2.4.1.68) precursor - human
N:Alternate names: alpha 1-6-FucT; GDP-L-fucose:N-acetyl-beta-D-glucosaminyl-6-alpha-L-f
C:Species: Homo sapiens (man)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000
C:Accession: JC5432; PC4322
R: Yanagidani, S.; Dozumi, N.; Ihara, Y.; Miyoshi, E.; Yamaguchi, N.; Taniguchi, N.
J. Biochem. 121, 626-632, 1997
A:Title: Purification and cDNA cloning of GDP-L-Fuc:N-acetyl-beta-D-glucosaminide:alpha
A:Reference number: JC5432; MUID:97279058
A:Accession: JC5432
A:Molecule type: mRNA
A:Residues: 1-575 <YAN>
A:Cross-references: DBJ:DB9289; NID:92055306; PIDN:BA19764.1; PID:g2055307
A:Accession: PC4322
A:Molecule type: Protein
A:Residues: 68-87/352-376/419-432 <YAN2>
C:Comment: This enzyme catalyzes the transfer of fucose from GDP-fucopyranoside to aspar
C:Genetics:
A:Gene: GDB:FUT8

A:Cross-references: GDB:9786294; OMIM:602589
A:Map position: 14q23-14q23
C:Superfamily: human glycoprotein 6-alpha-L-fucosyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-575/Product: glycoprotein 6-alpha-L-fucosyltransferase #status predicted <MAT>
Query Match 68.8%; Score 44; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 RRRPPYLP 10
Db 299 RRRPPYLP 305
RESULT 8
T29299
hypothetical protein C50F7.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T29299
R:Johnson, D.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C50F7.
A:Reference number: Z20601
A:Accession: T29299
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-589 <JON>
A:Cross-references: EMBL:U41557; PIDN:AAA83307.1; CESP:C50F7.2
C:Genetics:
A:Gene: CESP:C50F7.2
A:Introns: 12/2
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
Query Match 65.6%; Score 42; DB 2; Length 589;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 RRRPPYLP 11
Db 105 RRRPPYLP 114
RESULT 9
S43566
R01H10.8 protein (clone R01H10) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 17-Apr-1998
C:Accession: S43566
R:Lightning, J.
submitted to the EMBL Data Library, March 1994
A:Reference number: S43563
A:Accession: S43566
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-740 <LIG>
A:Cross-references: EMBL:Z31590
C:Genetics:
A:Introns: 23/1; 62/1; 116/3; 147/3; 260/3; 300/3; 431/2; 530/1; 595/2
C:Superfamily: SAM homology
F:48-115/Domain: SAM homology <SAM>
Query Match 65.6%; Score 42; DB 2; Length 740;
Best Local Similarity 63.6%; Pred. No. 79;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 RRRPPYLP 11

Db 368 KRPRPPQLPK 378

RESULT 10

hypothetical protein At2g22000 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84607
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 Euse, D.; Nleman, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayan, L.; Tallon, L.;
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: G84607
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-104 <STO>
 A:Cross-references: GB:AE002093; NID:g4417292; PIDN:AMD20417.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g22000
 A:Map position: 2

Query Match
 Best Local Similarity 64.1%; Score 41; DB 2; Length 104;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPP 7
 | | | | |
 Db 87 RRRPP 93

RESULT 11

WMBEXE

infected cell protein ICP34.5 - human herpesvirus 2 (strain HG52)
 N:Alternate names: RL1 protein
 C:Species: human herpesvirus 2
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
 R:McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.
 J. Gen. Virol. 72, 3057-3075, 1991
 A:Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
 A:Reference number: J01494; MUID:92113549
 A:Accession: J01502
 A:Molecule type: DNA
 A:Residues: 1-261 <MCG>
 A:Cross-references: GB:D10471; DDBJ:D01128; NID:g221784; PIDN:BAA23428.1; PID:g2626943
 C:Genetics:
 A:Gene: RL1
 A:Insertions: 165/3
 C:Superfamily: herpesvirus infected cell protein ICP34.5
 C:Keywords: tandem repeat
 F:3-12/Region: 5-residue repeats (R-R-R-G-P)
 F:16-31/Region: 8-residue repeats (P-R-P-G-A-P-A-V)

Query Match
 Best Local Similarity 64.1%; Score 41; DB 1; Length 261;
 Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 RRRPP-PYLR 11
 | | | | |
 Db 13 RRRPPGAPVR 25

RESULT 12

D96621
 auxin response factor 1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96621

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpspiel, N.A.; Kaul, S.; White, O.; Alon
 angen, N.F.; Hughes, B.; Huiztar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: D96621
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-665 <STO>
 A:Cross-references: GB:AE005173; NID:g5080809; PIDN:AAD39318.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F23H1.7
 A:Map position: 1

Query Match
 Best Local Similarity 64.1%; Score 41; DB 2; Length 665;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPPYL 10
 | | | | |
 Db 372 KRPRPPGLP 380

RESULT 13

T03861
 glycine-rich protein 15 - common tobacco (fragment)
 N:Alternate names: probable pollen wall protein
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jan-2000
 C:Accession: T03861
 R:Fuerstenberg, S.L.; Buccigaglia, P.A.; Smith, A.G.
 submitted to the EMBL Data Library, October 1996
 A:Description: Molecular characterization of an anther-specific gene from tobacco sho
 A:Reference number: Z15123
 A:Accession: T03861
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-129 <FUN>
 A:Cross-references: EMBL:U74325; NID:g1658178; PIDN:AMB18261.1; PID:g1658179
 C:Genetics:
 A:Gene: GRP15
 C:Superfamily: Phaeocelus glycine-rich cell wall protein 1.8

Query Match
 Best Local Similarity 62.5%; Score 40; DB 2; Length 129;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPPYL 9
 | | | | |
 Db 28 RRRPPYL 35

RESULT 14

A70577
 hypothetical protein RV2133c - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: A70577
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987
A:Accession: A70577
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-262 <COL>
A:Cross-references: GB:595388; GB:AL123456; NID:93261759; PIDN:CAR08660.1; PID:92104345
C:Genetics:
A:Gene: RV2133C

Query Match 62.5%; Score 40; DB 2; Length 262;
Best Local Similarity 87.5%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRPRLP 10
||| |||
Db 116 RPRGYP 123

RESULT 15
B23971
monophenol monooxygenase (EC 1.14.18.1) - Streptomyces antibioticus
N:Alternate names: tyrosinase
C:Species: Streptomyces antibioticus
C>Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 22-Oct-1999
C:Accession: B23971
R:Bernan, V.; Filipula, D.; Herber, W.; Bibb, M.; Katz, E.
Gene 37, 101-110, 1985
A:Title: The nucleotide sequence of the tyrosinase gene from Streptomyces antibioticus
A:Reference number: A91531; MUID:86031341
A:Accession: B23971
A:Molecule type: DNA
A:Residues: 1-273 <BER>
A:Cross-references: GB:M11582; NID:9153522; PIDN:AAA88571.1; PID:9153523
C:Keywords: oxidoreductase

Query Match 62.5%; Score 40; DB 2; Length 273;
Best Local Similarity 70.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRRPYP 10
||| |||
Db 228 RRRSPYP 237

RESULT 16
J01663
hybrid proline-rich protein - maize
C:Species: Zea mays (maize)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Sep-1999
C:Accession: J01663
R:Jose-Estanyol, M.; Ruiz-Avila, L.; Puigdomenech, P.
Plant Cell 4, 413-423, 1992
A:Title: A maize embryo-specific gene encodes a proline-rich and hydrophobic protein.
A:Reference number: J01663; MUID:92361259
A:Accession: J01663
A:Molecule type: DNA
A:Residues: 1-301 <ROS>
A:Cross-references: EMBL:X60432; NID:9433706; PIDN:CAA42959.1; PID:9433707
A:Experimental source: strain W64A
C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 62.5%; Score 40; DB 2; Length 301;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRPRLP 10
||| |||
Db 83 RSPYP 90

RESULT 17
T01316
epoxide hydrolase homolog T14P8.15 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Nov-1999
C:Accession: T01316
R:Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of A. thaliana T14P8.
A:Reference number: Z14290
A:Accession: T01316
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-324 <KAL>
A:Cross-references: EMBL:AF069298; NID:93193282; PID:93193297
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 155/3; 235/2
A:Note: T14P8.15
C:Superfamily: tropinesterase

Query Match 62.5%; Score 40; DB 2; Length 324;
Best Local Similarity 60.0%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRPRLP 11
||| |||
Db 180 RMRPRLP 189

RESULT 18
160916
HNF-3/forkhead homolog-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I60916
R:Clevidence, D.E.; Overdier, D.G.; Tao, W.; Qian, X.; Panl, L.; Lal, E.; Costa, R.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 3948-3952, 1993
A:Title: Identification of nine tissue-specific transcription factors of the hepatocyt
A:Reference number: A47450; MUID:93248207
A:Accession: I60916
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-439 <RES>
A:Cross-references: GB:U13201; NID:951067; PIDN:AA474561.1; PID:9550513
C:Genetics:
A:Gene: HNF-1
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:100-192/Domain: fork head DNA-binding domain homology <FHD>

Query Match 62.5%; Score 40; DB 2; Length 439;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRPRLP 8
||| |||
Db 97 RRPRLP 103

RESULT 19
S63033
hypothetical protein YNL094w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N2219
C:Species: Saccharomyces cerevisiae
C>Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 17-Mar-1999
C:Accession: S63033; S52731; S69995
R:Garcia-Cantalejo, J.M.; Boskovic, J.; Jimenez, A.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63030

A:Accession: S63033
 A:Molecule type: DNA
 A:Residues: 1-587 <GAR>
 A:Cross-references: EMBL:Z71370; NID:g1301999; PID:e239791; PID:g1302000; MIPS:YNI094w
 R: Garcia-Cantalejo, J.M.; Boskovic, J.; Jimenez, A.
 Submitted to the EMBL Data Library, March 1995
 A:Reference number: S52729
 A:Accession: S52731
 A:Molecule type: DNA
 A:Residues: 182-587 <GAW>
 A:Cross-references: EMBL:X85811; NID:g758291; PID:g758294
 R: Garcia-Cantalejo, J.M.; Boskovic, J.; Jimenez, A.
 A:Title: Sequence analysis of a 14.2 kb fragment of *Saccharomyces cerevisiae* chromosome
 A:Reference number: S69993; MUID:96367601
 A:Accession: S69995
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-587 <GAR>
 A:Cross-references: EMBL:X85811
 C:Genetics:
 A:Map position: 14L
 A:Note: YNI094w

Query Match
 Best Local Similarity 62.5%; Score 40; DB 2; Length 587;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPVLP 10
 |||||
 Db 466 RRRPPVLP 475

RESULT 20
 G83081
 conserved hypothetical protein PA4517 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83081
 R:Stover, C.K.; Pham, X.Q.; Eyrin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.D.; B.
 ; Lory, S.; Olson, M.V.
 Nature 406: 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337
 A:Accession: G83081
 A:Molecule type: DNA
 A>Status: preliminary
 A:Residues: 1-600 <STO>
 A:Cross-references: GB:AE004865; GB:AE004091; NID:g9950752; PIDN:AA07905.1; GSPDB:GN001
 C:Genetics:
 A:Experimental source: strain PA01
 A:Gene: PA4517

Query Match
 Best Local Similarity 62.5%; Score 40; DB 2; Length 600;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPPVLP 11
 |||||
 Db 141 RRRPPVLP 150

RESULT 21
 JN0673
 ubiquitin-like fusion protein Anla - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 08-Dec-2000
 R:Linmen, J.M.; Bailey, C.P.; Weeks, D.L.

Gene 128, 181-188, 1993
 A:Title: Two related localized mRNAs from *Xenopus laevis* encode ubiquitin-like fusion
 A:Reference number: JN0673; MUID:93292985
 A:Accession: JN0673
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-693 <LIN>
 C:Genetics:
 A:Gene: Anla
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
 C:Keywords: zinc finger
 F:28-103/Region: ubiquitin-like protein
 F:28-103/Domain: ubiquitin homology <UBH>
 F:625-693/Region: zinc finger

Query Match
 Best Local Similarity 62.5%; Score 40; DB 2; Length 693;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPVLP 11
 |||||
 Db 229 KRRPPVLP 237

RESULT 22
 JN0674
 ubiquitin-like fusion protein Anlb - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 08-Dec-2000
 C:Accession: JN0674
 R:Linmen, J.M.; Bailey, C.P.; Weeks, D.L.
 Gene 128, 181-188, 1993
 A:Title: Two related localized mRNAs from *Xenopus laevis* encode ubiquitin-like fusion
 A:Reference number: JN0673; MUID:93292985
 A:Accession: JN0674
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-701 <LIN>
 A:Cross-references: GB:L08475; NID:g214867; PIDN:AAA49979.1; PID:g214868
 C:Genetics:
 A:Gene: Anlb
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
 C:Keywords: zinc finger
 F:28-103/Region: ubiquitin-like protein
 F:28-103/Domain: ubiquitin homology <UBH>
 F:633-701/Region: zinc finger

Query Match
 Best Local Similarity 62.5%; Score 40; DB 2; Length 701;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPVLP 11
 |||||
 Db 228 KRRPPVLP 236

RESULT 23
 T19140
 hypothetical protein C09G5.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 R:Palmer, S.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: Z19080
 A:Accession: T19140
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-963 <WIL>
 A:Cross-references: EMBL:Z46791; PIDN:CAA86755.1; GSPDB:GN00020; CESP:C09G5.6
 A:Experimental source: clone C09G5

A:Accession: G82780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <SIM>
 A:Cross-references: GB:AE003909; GB:AE003849; NID:g9105513; PIDN:AAFB3459.1; GSPDB:GN001
 A:Experimental source: strain 945c
 R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H.
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 A:Submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B.
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; de Sa, R.G.; Santelli, R.V.; Sawasak
 M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0649

Query Match 60.9%; Score 39; DB 2; Length 224;
 Best Local Similarity 62.5%; Pred. No. 66;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 RRRPPYPLR 9
 Db 213 RRRPPYPLR 220

RESULT 29
 T49750
 related to DNA-directed RNA polymerase II largest chain [imported] - Neurospora crassa
 N:Alternate names: protein B24B19.220
 C:Species: Neurospora crassa
 C>Date: 02-Jun-2000 #sequence-revision 02-Jun-2000 #text-change 02-Jun-2000
 C:Accession: T49750
 R:Schulze, U.; Allyn, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: 225022
 A:Accession: T49750
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <SCH>
 A:Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.220
 A:Experimental source: BAC clone B24B19; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B24B19.220
 A:Map position: 6

Query Match 60.9%; Score 39; DB 2; Length 333;
 Best Local Similarity 70.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 RRRPPYPLR 11
 Db 207 RRRPPYPLR 216

RESULT 30
 B82634
 Chreonline dehydratase catabolic XFI819 [imported] - Xylella fastidiosa (strain 945c)
 C:Species: Xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text-change 20-Aug-2000
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82634
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-380 <SIM>
 A:Cross-references: GB:AE004003; GB:AE003849; NID:g9106894; PIDN:AAFB4625.1; GSPDB:GN
 A:Experimental source: strain 945c
 R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 A:Submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B.
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; de Sa, R.G.; Santelli, R.V.; Sawasak
 M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XFI819

Query Match 60.9%; Score 39; DB 2; Length 380;
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 3 RRRPPYPLR 11
 Db 354 RRRPPYPLR 362

RESULT 31
 G85957
 unknown protein encoded by ISE8 [imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 31-Mar-2001
 C:Accession: G85957
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamouls, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: G85957
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-467 <STO>
 A:Cross-references: GB:AE005174; NID:g12517550; PIDN:AAG58123.1; GSPDB:GN00145; UMG
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z4340

Query Match 60.9%; Score 39; DB 2; Length 467;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 RRRPPYPLR 11
 Db 50 RRRPPYPLR 59

RESULT 32
 T33739
 hypothetical protein K08B4.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence-revision 29-Oct-1999 #text-change 29-Oct-1999
 R:Rohlfing, T.; Murray, J.; Antonio, B.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid K08B4.

A:Reference number: 221395
A:Accession: T33739
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-488 <ROH>
A:Cross-references: EMBL:AF100663; PIDN:AAC68982.1; GSPDB:GN00022; CESP:K08B4.5
C:Genetics:
A:Experimental source: strain Bristol N2; clone K08B4
A:Gene: CESP:K08B4.5
A:Map position: 4
A:Introns: 222/1; 318/2; 342/2; 399/3; 409/3; 440/2

Query Match
Best Local Similarity 60.9%; Score 39; DB 2; Length 488;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPPPY 8
|||||
Db 345 RRRPNP 352

RESULT 33
C96817
F9K20.15 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R:Accession: C96817
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: C96817
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-505 <STO>
A:Cross-references: GB:AE005173; NID:g3834315; PIDN:AAC83031.1; GSPDB:GN00141
C:Genetics:
A:Gene: F9K20.15
A:Map position: 1

Query Match
Best Local Similarity 60.9%; Score 39; DB 2; Length 505;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPPPY 10
|||||
Db 43 QRRPNY 51

RESULT 34
F85605
unknown in ISEC8 [imported] - Escherichia coli (strain 0157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: F85605
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoukis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: F85605
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <STO>

A:Cross-references: GB:AE005174; NID:g12513963; PIDN:AAG55306.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1161

Query Match
Best Local Similarity 60.9%; Score 39; DB 2; Length 537;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPPY 11
|||||
Db 120 RRRPAP 129

RESULT 35
G85656
unknown in ISEC8 [imported] - Escherichia coli (strain 0157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: G85656
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoukis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: G85656
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <STO>
A:Cross-references: GB:AE005174; NID:g12514483; PIDN:AAG55715.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1600

Query Match
Best Local Similarity 60.9%; Score 39; DB 2; Length 537;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPPY 11
|||||
Db 120 RRRPAP 129

RESULT 36
T27512
hypothetical protein ZC302.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27512
R:Kelly, P.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z20380
A:Accession: T27512
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-772 <MIL>
A:Cross-references: EMBL:Z73978; PIDN:CAA98292.1; GSPDB:GN00023; CESP:ZC302.1
A:Experimental source: clone ZC302
C:Genetics:
A:Gene: CESP:ZC302.1
A:Map position: 5
A:Introns: 43/2; 70/3; 94/2; 129/2; 393/3; 548/3; 598/3; 642/3

Query Match
Best Local Similarity 60.9%; Score 39; DB 2; Length 772;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPPY 10
|||||
Db 462 KRRRPP 471

```

RESULT 37
S06124
retinoic acid receptor delta - eastern newt
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Aug-1999
C:Accession: S06124; S09180
R:Ragsdale Jr., C.W.; Petkovich, M.; Gates, P.B.; Chambon, P.; Brookes, J.P.
Nature 341, 654-657, 1989
A:Title: Identification of a novel retinoic acid receptor in regenerative tissues of the
A:Reference number: S06123; MUID:90015190
A:Accession: S06124
A:Molecule type: mRNA
A:Residues: 1-454 <RAG>
A:Cross-references: EMBL:X17586; NID:664134; PIDN:CAA35603.1; PID:g64135
R:Giguere, V.; Ong, E.S.; Evans, R.M.; Tabin, C.J.
Nature 337, 566-569, 1989
A:Title: Spatial and temporal expression of the retinoic acid receptor in the regenerati
A:Reference number: S02758; MUID:89127522
A:Accession: S09180
A:Molecule type: mRNA
A:Residues: 215-229,'D',231-251,'S',253-445,'A',447-454 <GIG>
C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F:90-341/Domain: erba transforming protein homology <ERBA>

Query Match
Best Local Similarity 60.2%; Score 38.5; DB 2; Length 454;
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 1 RRRPRPYL-PR 11
Db 369 RRRPRPYL-PR 380

RESULT 38
S75341
hypothetical protein slr2010 - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S75341
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <KAN>
A:Cross-references: EMBL:D90904; GB:AB001339; NID:g1652225; PIDN:BA017255.1; PID:d101798
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match
Best Local Similarity 59.4%; Score 38; DB 2; Length 130;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPRPYL-PR 11
Db 38 RRRPRPYL-PR 47

RESULT 39
S09612
hypothetical XrepBE protein - human
C:Species: Homo sapiens (man)
C:Date: 12-Mar-1999 #sequence_revision 12-Mar-1999 #text_change 12-Mar-1999
C:Accession: S09612

```

```

R:Riley, D.E.
Nucleic Acids Res. 17, 2361, 1989
A:Title: Nucleotide sequence of the Plasmid stimulating portion of Xrep.
A:Reference number: S09612; MUID:89202053
A:Accession: S09612
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-139 <RIL>
A:Cross-references: EMBL:X13896; NID:g37998
C:Comment: This sequence is not thought to be translated.
C:Genetics:
A:Gene: Xrep

Query Match
Best Local Similarity 59.4%; Score 38; DB 4; Length 139;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 RRRPRPYL-PR 11
Db 60 RRRPRPYL-PR 68

RESULT 40
A03863
hypothetical protein F-215 - human adenovirus 2
C:Species: Mastadenovirus h2 (human adenovirus 2)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 31-Dec-1993
C:Accession: A43046; A43047; A03863
R:Gingras, T.R.; Sciaky, D.; Gellinas, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; B
J. Biol. Chem. 257, 13475-13491, 1982
A:Title: Nucleotide sequences from the adenovirus-2 genome.
A:Reference number: A92351; MUID:83056843
A:Accession: A43046
A:Molecule type: DNA
A:Residues: 1-215 <GIN>
R:Alestrom, P.; Akusjarvi, G.; Pettersson, M.; Pettersson, U.
J. Biol. Chem. 257, 13492-13498, 1982
A:Title: DNA sequence analysis of the region encoding the terminal protein and the hy
A:Reference number: A92352; MUID:83056844
A:Accession: A43047
A:Molecule type: DNA
A:Residues: 1-215 <ALE>

Query Match
Best Local Similarity 59.4%; Score 38; DB 2; Length 215;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPRPYL-PR 9
Db 135 RRRPRPYL-PR 143

RESULT 41
T43472
hypothetical protein DKFZp434F1728.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T43472
R:Blocker, H.; Boeher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z22513
A:Accession: T43472
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-237 <AAA>
A:Cross-references: EMBL:AL133615
A:Experimental source: adult testis; clone DKFZp434F1728
C:Genetics:
A:Note: DKFZp434F1728.1

```

Query Match 59.4%; Score 38; DB 2; Length 237;
 Best Local Similarity 85.7%; Pred. No. 98;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPY 7
 |||||
 DB 224 RRRPKPP 230

RESULT 42

S28827
 chlorophyll a/b-binding protein type I - common tobacco
 M:Alternate names: light-harvesting chlorophyll a/b-binding protein
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 26-Aug-1999
 C:Accession: S28827
 R:Palomares, R.; Herrmann, R.G.; Oelmueller, R.
 J. Photochem. Photobiol. B 11, 151-162, 1991
 A:Title: Different blue-light requirement for the accumulation of transcripts from nucle
 A:Reference number: S2735; MUID:92122059
 A:Accession: S28827
 A:Molecule type: mRNA
 A:Residues: 1-245 <PAL>
 A:Cross-references: EMBL:X64198; NID:9493722; PIDN:CAA45523.1; PID:9493723
 A:Note: the authors translated the codon CTT for residue 240 as Pro
 C:Superfamily: chlorophyll a/b-binding protein
 C:Keywords: chlorophyll; chloroplast; light-harvesting complex; photosystem I; thylakoid

Query Match 59.4%; Score 38; DB 2; Length 245;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 RRRPPY 9
 :|||||
 DB 53 QRRPPYL 59

RESULT 43

T30959
 hypothetical protein F42A10.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C:Accession: T30959
 R:Latreille, P.
 submitted to the EMBL Data Library, August 1999
 A:Description: The sequence of C. elegans cosmid F42A10.
 A:Reference number: Z20946
 A:Accession: T30959
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-272 <LAT>
 A:Cross-references: EMBL:U10414; PIDN:AAA19076.1
 A:Experimental source: strain Bristol N2; clone F42A10
 C:Genetics:
 A:Map position: III
 A:Introns: 75/2; 229/3
 A:Note: F42A10.5

Query Match 59.4%; Score 38; DB 2; Length 272;
 Best Local Similarity 70.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRRPPYLP 11
 ||| |||
 DB 143 RRGAPYRPR 152

RESULT 44

T35646
 probable thiamin monophosphate kinase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Feb-2000
 C:Accession: T35646
 R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
 submitted to the EMBL Data Library, December 1998

A:Reference number: Z21585
 A:Accession: T35646
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-322 <MUR>
 A:Cross-references: EMBL:AL034447; PIDN:CAA22405.1; GSPDB:GN00070; SCOEDB:SC7A1.06
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC7A1.06
 C:Superfamily: hydrogenase expression/formation protein hye

Query Match 59.4%; Score 38; DB 2; Length 322;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPY 8
 ||| |||
 DB 192 RRRPPY 198

RESULT 45

A47234
 homeobox protein H6 - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 17-Oct-1997
 C:Accession: A47234
 R:Stadler, H.S.; Padanilam, B.J.; Buetow, K.; Murray, J.C.; Solursh, M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 11579-11583, 1992
 A:Title: Identification and genetic mapping of a homeobox gene to the 4p16.1 region o
 A:Reference number: A47234; MUID:93087572
 A:Accession: A47234
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-373 <STA>
 A:Experimental source: embryo craniofacial region
 A:Note: sequence extracted from NCBI backbone (NCBIN:119953, NCBI:119955)
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:201-257/Domain: homeobox homology <HOX>

Query Match 59.4%; Score 38; DB 2; Length 373;
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 RRRPPYLP 11
 ||||| ||
 DB 288 RRRPPYLP 296

Search completed: September 24, 2001, 10:06:05
 Job time: 45 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 24, 2001, 10:09:04 : Search time 16.18 Seconds
(Without alignments)
23.289 Million cell updates/sec

Title: US-09-276-868-4

Sequence: 1 RRRPRPYLPR 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	172	PR39_PIG	P80054 sus scrofa
2	47	73.4	190	BCF7_BOVIN	P19661 bos taurus
3	45	70.3	372	DBPA_HUMAN	P16989 homo sapien
4	44	68.8	107	CORT_BACSU	P18683 bacillus su
5	41	64.1	261	RL1_HSV2H	P28283 herpes simp
6	41	64.1	625	DUS8_HUMAN	Q13202 homo sapien
7	40	62.5	151	RNB_HSV2H	P99479 herpes simp
8	40	62.5	272	TYRO_STRAT	P07524 streptomyc
9	40	62.5	333	HME1_CHICK	O05916 gallus gall
10	40	62.5	439	HPH1_RAT	O63244 rattus norv
11	40	62.5	587	YN04_YEAST	P33933 saccharomyc
12	40	62.5	963	YQ36_CAEEL	O09457 capnorbadi
13	39	60.9	148	YL12_EBV	P03139 epstein-bar
14	39	60.9	148	YL12_EBV	O07285 epstein-bar
15	39	60.9	180	XG_HUMAN	P55808 homo sapien
16	39	60.9	190	BCT7_SHEEP	P50415 ovis aries
17	39	60.9	233	MTRP_HUMAN	Q15012 homo sapien
18	39	60.9	233	MTRP_MOUSE	O60961 mus musculu
19	39	60.9	393	CIM4_HUMAN	Q09y98 homo sapien
20	38.5	60.2	505	RRG_NORVI	P18516 notophthalp
21	38	59.4	161	RNB_HSV1M	P69568 herpes simp
22	38	59.4	215	Y215_ADE02	P03291 human adeno
23	38	59.4	340	ADHA_RHIME	O31186 trizobolun m
24	38	59.4	508	MKR1_YEAST	P32180 saccharomyc
25	38	59.4	520	CP84_ARATH	Q42600 arabidopsi
26	38	59.4	845	PPSA_AERPE	O9yec5 aeropyrum p
27	38	59.4	846	IRSL_HCMVA	P09375 human cytom
28	37.5	58.6	443	RRG2_HUMAN	P22933 homo sapien
29	37.5	58.6	454	RRG1_HUMAN	P13631 homo sapien
30	37.5	58.6	543	PKAA_STRCO	P54739 streptomyc
31	37	57.8	498	ARAA_BACSU	P45523 bacillus su
32	37	57.8	581	FKBX_MOUSE	O61576 mus musculu
33	37	57.8	613	SG2_BOVIN	P20616 bos taurus

34	37	57.8	904	1	VGLB_HSV11	P10211 herpes simp
35	37	57.8	1016	1	PCR2_SCHPO	P36583 schistosach
36	37	57.8	1047	1	GTPA_HUMAN	P20936 homo sapien
37	37	57.8	1342	1	ERB3_HUMAN	P21860 homo sapien
38	37	57.8	2414	1	P300_HUMAN	Q09472 homo sapien
39	36.5	57.0	442	1	RRG2_XENLA	P28699 xenopus lae
40	36.5	57.0	476	1	RRG1_XENLA	P51127 xenopus lae
41	36	56.2	80	1	SSS2_SCYCA	P11020 scylliorhinu
42	36	56.2	80	1	GT22_ARATH	O92vc3 arabidopsi
43	36	56.2	221	1	ICP3_HSV1N	P37319 herpes simp
44	36	56.2	245	1	ICP3_HSV1I	P36313 herpes simp
45	36	56.2	252	1	ICP3_HSV1D	P37318 herpes simp

ALIGNMENTS

RESULT	ID	PR39_PIG	STANDARD	PRT	172 AA.
AC	P80054	Q9TR84			
DT	01-MAR-1992	(Rel. 21, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	ANTIBACTERIAL PROTEIN PR-39 PRECURSOR.				
GN	PR39.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=95350216; PubMed=7624374;				
RA	Gudmundsson G.H., Magnusson K.P., Chowdhary B.P., Johansson M.,				
RA	Andersson L., Boman H.G.;				
RT	*Structure of the gene for porcine peptide antibiotic PR-39, a				
RT	cathelin gene family member: comparative mapping of the locus for the				
RT	human peptide antibiotic FALL-39.;				
RL	Proc. Natl. Acad. Sci. U.S.A. 92:7085-7089(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RX	MEDLINE=94071853; PubMed=8250863;				
RA	Storici P., Zanetti M.;				
RT	*A cDNA derived from pig bone marrow cells predicts a sequence				
RT	identical to the intestinal antibacterial peptide PR-39.;				
RL	Biochem. Biophys. Res. Commun. 196:1058-1065(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=96105365; PubMed=7496526;				
RA	Zhao C., Ganz T., Lehrer R.I.;				
RT	*Structures of genes for two cathelin-associated antimicrobial				
RT	peptides: prophenin-2 and PR-39.;				
RL	FEBS Lett. 376:130-134(1995).				
RN	[4]				
RP	SEQUENCE OF 131-169.				
RC	TISSUE=Intestine;				
RX	MEDLINE=92111534; PubMed=1765098;				
RA	Agdeher B., Lee J.-Y., Bergman T., Carlquist M., Boman H.G.,				
RT	Mutt V., Joernvall H.;				
RT	*Amino acid sequence of PR-39. Isolation from pig intestine of a new				
RT	member of the family of proline-arginine-rich antibacterial				
RT	peptides.;				
RL	Eur. J. Biochem. 202:849-854(1991).				
RN	[5]				
RP	SEQUENCE OF 131-164, AND FUNCTION.				
RC	TISSUE=Neutrophils;				
RX	MEDLINE=95088504; PubMed=7996056;				
RA	Shi J., Ross C.R., Chengappa M.M., Blecha F.;				
RT	*Identification of a proline-arginine-rich antibacterial peptide from				
RT	neutrophils that is analogous to PR-39, an antibacterial peptide from				
RT	the small intestine.;				

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (FORM 3).
 RX MEDLINE-89211987; PubMed-2977358;
 RA Sakura H., Maekawa T., Imamoto F., Yasuda K., Ishii S.;
 RT "Two human genes isolated by a novel method encode DNA-binding
 RL proteins containing a common region of homology.";
 RN Gene 73:499-507(1988).
 RN [2]
 RP SEQUENCE FROM N.A. (FORMS 2 AND 3).
 RX TISSUE-Placenta;
 MEDLINE-9554705; PubMed-7628487;
 RA Kudo S., Mattei M.-C., Fukuda M.;
 RT "Characterization of the gene for dbpA, a family member of the
 RL nucleic-acid-binding proteins containing a cold-shock domain";
 RN Eur. J. Biochem. 231:72-82(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (FORM 1).
 RX MEDLINE-96279731; PubMed-8710501;
 RA Coles L.S., Diamond P., Occhiodoro F., Vadas M.A., Shannon M.F.;
 RT "Cold shock domain proteins repress transcription from the GM-CSF
 RL promoter";
 RN Nucleic Acids Res. 24:2311-2317(1996).
 CC -1- FUNCTION: BINDS TO THE GM-CSF PROMOTER. SEEMS TO ACT AS A
 CC REPRESSOR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2
 CC AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE AND HEART.
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M24069; AAA55749.1; ALT_INIT.
 CC EMBL: L29071; AAA79243.1; -.
 CC EMBL: L29064; AAA79243.1; JOINED.
 CC EMBL: L29065; AAA79243.1; JOINED.
 CC EMBL: L29066; AAA79243.1; JOINED.
 CC EMBL: L29067; AAA79243.1; JOINED.
 CC EMBL: L29068; AAA79243.1; JOINED.
 CC EMBL: L29069; AAA79243.1; JOINED.
 CC EMBL: L29070; AAA79243.1; JOINED.
 CC EMBL: X95325; CAA64631.1; -.
 CC PIR: PS0014; PS0014.
 CC HSSP: P15277; IMJC.
 CC TRANSFAC: T00185; -.
 CC MIM: 603437; -.
 CC InterPro: IPR002059; -.
 CC Pfam: PF00313; CSD, 1.
 CC PRINTS: PR00050; COLDSHOCK.
 CC PROSITE: PS00352; COLD_SHOCK, 1.
 CC Transcription regulation; Repressor; DNA-binding; Nuclear protein;
 CC Alternative splicing.
 CC KW DOMAIN 93 157 CSD.
 CC FT VARSPIC 192 260 MISSING (IN ISOFORM 2).
 CC FT VARSPIC 340 372 RPPNPSDDGKFAKAGEALTENPAPPTQSSAE -> PSS
 CC FT CONFLICT 75 75 A -> T (IN REF. 3).
 CC FT SEQUENCE 372 AA; 40060 MW; A52402AC3P6385CC CRC64;
 CC -----
 CC Query Match 70.3%; Score 45; DB 1; Length 372;
 CC Best Local Similarity 80.0%; Pred. No. 5.6;
 CC Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPPYLP 10
 DB 336 RRRPPPNAP 345

 CC RESULT 4
 CC COTT_BACSU STANDARD; PRT; 107 AA.
 AC P11863;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE SPORE COAT PROTEIN T PRECURSOR.
 GN COTT.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 / JH642;
 RX MEDLINE-89313296; PubMed-2546006;
 RA Aronson A.I., Song H.Y., Bourne N.;
 RT "Gene structure and precursor processing of a novel Bacillus subtilis
 RL spore coat protein";
 RN Mol. Microbiol. 3:437-444(1989).
 CC -1- FUNCTION: POSSIBLY PROTECTION OF SPORE AND PROBABLY PLAYS
 CC SOME ROLE IN GERMINATION.
 CC -1- SUBCELLULAR LOCATION: OUTER SURFACE OF ENDOSPORE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X13740; CAA32004.1; -.
 CC EMBL: Z99110; CAB13066.1; -.
 CC PIR: S04835; S04835.
 CC Subtilist; Bg10495; cott.
 CC Sporulation; Signal.
 CC FT SIGNAL 1 44
 CC FT CHAIN 45 107 SPORE COAT PROTEIN T.
 CC FT SEQUENCE 107 AA; 12992 MW; ADIF66F0C4CE29A3 CRC64;
 CC -----
 CC Query Match 68.8%; Score 44; DB 1; Length 107;
 CC Best Local Similarity 87.5%; Pred. No. 2.2;
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC OY 3 RRRPPPYLP 10
 CC DB 82 RRRPPPYLP 89
 CC -----
 CC RESULT 5
 CC RLI_HSV2H STANDARD; PRT; 261 AA.
 ID RLI_HSV2H
 AC P28283;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE NEUROVIRULENCE FACTOR (ICP34.5).
 GN RLI.
 OS Herpes simplex virus (type 2 / strain HG52).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxId=10315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92113549; PubMed-1662697;

RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
 RT "Comparative sequence analysis of the long repeat regions and
 RT adjoining parts of the long unique regions in the genomes of herpes
 RT simplex viruses 1 and 2."
 RL J. Gen. Virol. 72:3057-3075(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10471; BA023428.1; -
 DR EMBL; 286099; CAB06759.1; -
 DR EMBL; 286099; CAB06706.1; -
 DR PIR; J01502; WMBEXE.
 KW Repeat.
 FT DOMAIN 3 12 2 X 5 AA TANDEM REPEATS OF R-R-R-G-P.
 FT REPEAT 3 7
 FT REPEAT 8 12
 FT DOMAIN 16 31 2 X 8 AA TANDEM REPEATS OF P-R-P-G-A-P-A-
 FT REPEAT 16 23 V.
 FT REPEAT 24 31
 FT SEQUENCE 261 AA; 27908 MW; 4BBD13AF3D906D71 CRC64;
 SQ
 Query Match
 Best Local Similarity 64.1%; Score 41; DB 1; Length 261;
 Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;
 QY 1 RRRPRP-PYLPR 11
 DB 13 RRRPRGAPAVPR 25
 RESULT 6
 DUS8_HUMAN STANDARD; PRT; 625 AA.
 ID DUS8_HUMAN
 AC Q13202;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH-5).
 GN DUSP8 OR VH5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Fetal brain;
 RX MEDLINE=96009533; PubMed=7561881;
 RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
 RT "HVH-5: a protein tyrosine phosphatase abundant in brain that
 RT inactivates mitogen-activated protein kinase."
 RL J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
 CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U27193; AAA83151.1; -
 DR MIM; 602038;
 DR InterPro; IPR000340; -
 DR InterPro; IPR000387; -
 DR InterPro; IPR002965; -
 DR Pfam; PF00782; DSPC; 1.
 DR PRINTS; PR01217; PRICHTXTNSN.
 DR PROSITE; PS00883; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 28 43 CH2 A DOMAIN.
 FT DOMAIN 117 132 CH2 B DOMAIN.
 FT DOMAIN 162 430 CATALYTIC.
 FT DOMAIN 310 550 PRO-RICH.
 FT ACT_SITE 246 246 BY SIMILARITY.
 FT SEQUENCE 625 AA; 65840 MW; DCBEA14487219666 CRC64;
 SQ
 Query Match
 Best Local Similarity 64.1%; Score 41; DB 1; Length 625;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRP 7
 DB 454 RRRPRP 460
 RESULT 7
 RNB_HSV2H STANDARD; PRT; 151 AA.
 ID RNB_HSV2H
 AC P89479;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE POTENTIAL RNA-BINDING PROTEIN.
 GN US11.
 OS Herpes simplex virus (type 2 / strain HG52).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: BINDS DNA AND RNA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z86099; CAB06719.1; -
 DR DNA-Binding; RNA-binding; Repeat; Nuclear protein.
 KW DNA-Binding; RNA-binding; Repeat; Nuclear protein.
 FT DOMAIN 90 146 11 X 6 AA TANDEM REPEATS.
 FT REPEAT 90 95 1.
 FT REPEAT 96 101 2.
 FT REPEAT 102 104 3.
 FT REPEAT 105 110 4.
 FT REPEAT 111 116 5.
 FT REPEAT 117 122 6.

```

FT REPEAT 123 128 7.
FT REPEAT 129 130 8.
FT REPEAT 130 134 9.
FT REPEAT 135 140 10.
FT REPEAT 141 146 11.
SQ SEQUENCE 151 AA; 16297 MW; FAB751F23C3DB6AE CRC64;

```

Query Match 62.5%; Score 40; DB 1; Length 151;
 Best Local Similarity 70.0%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY 2 RRRPPPYLP 11
    |||||:11
DB 127 RRRPPPYLP 136

```

```

RESULT 8
TYRO_STRAT ID TYRO_STRAT STANDARD; PRT: 272 AA.
AC P07524;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TYROSINASE (EC 1.14.18.1) (MONOPHENOL MONOOXYGENASE).
GN MEJC2 OR MEL.
OS Streptomyces antibioticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1890;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66031341; PubMed=3932128;
RA Bernan V., Filipula D., Herper W., Bibb M.J., Katz E.;
  "The nucleotide sequence of the tyrosinase gene from Streptomyces
  antibioticus and characterization of the gene product.";
  Gene 37:101-110(1985).
CC -1- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
  THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
  COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: L-TYROSINE + L-DOPA + O(2) = L-DOPA +
  DOPAQUINONE + H(2)O.
CC -1- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
DR EMBL; M11582; AAA8571.1; -
DR PIR; B23971; B23971.
DR InterPro; IPR002227; -
DR Pfam; PF00264; tyrosinase; 1.
DR PRINTS; PR00092; TYROSINASE; 1.
DR PROSITE; PS00497; TYROSINASE_1; 1.
DR PROSITE; PS00498; TYROSINASE_2; 1.
KW Melanin biosynthesis; Oxidoreductase; Monooxygenase; Copper.
FT INT_MET 0
FT METAL 37
FT METAL 37 COPPER A (BY SIMILARITY).
FT METAL 53 COPPER A (BY SIMILARITY).
FT METAL 62 COPPER A (BY SIMILARITY).
FT METAL 189 COPPER B (BY SIMILARITY).
FT METAL 193 COPPER B (BY SIMILARITY).
FT METAL 193 COPPER B (BY SIMILARITY).
FT METAL 215 COPPER B (BY SIMILARITY).
FT METAL 215 COPPER B (BY SIMILARITY).
SQ SEQUENCE 272 AA; 30608 MW; E10A9A802D9CDBA CRC64;

```

Query Match 62.5%; Score 40; DB 1; Length 272;
 Best Local Similarity 70.0%; Pred. No. 22;

```

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 RRRPPPYLP 10
    |||1111
DB 227 RRRPPPYLP 236

```

```

RESULT 9
HME1_CHICK ID HME1_CHICK STANDARD; PRT: 333 AA.
AC Q05916;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HOMEBOX PROTEIN ENGRAILED-1 (GG-EN-1).
GN EN1 OR EN-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93185339; PubMed=1363401;
RA Logan C., Hanks M.C., Noble-Topham S., Nallathath D.,
  Provart N.J., Joyner A.L.;
  "Cloning and sequence comparison of the mouse, human, and chicken
  engrailed genes reveal potential functional domains and regulatory
  regions.";
  Dev. Genet. 13:345-358(1992).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ENGRAILED FAMILY OF HOMEBOX PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC EMBL; L12694; AAA53435.1; -
CC EMBL; L12695; AAA53436.1; -
CC HSSP; P02836; 1HDD.
DR TRANSFAC; T02014; -
DR InterPro; IPR000747; -
DR InterPro; IPR001356; -
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00026; ENGRAILED.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR PROSITE; PS00033; ENGRAILED; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 22 29 POLY-GLY.
FT DOMAIN 36 65 PRO-RICH.
FT DOMAIN 59 65 POLY-PRO.
FT DOMAIN 97 104 POLY-GLY.
FT DNA_BIND 244 303 HOMEBOX.
SQ SEQUENCE 333 AA; 34515 MW; D0F1B1F917E1FBAD CRC64;

```

Query Match 62.5%; Score 40; DB 1; Length 333;
 Best Local Similarity 72.7%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 RRRPPPYLP 11
    |||||111
DB 56 RRRPPPYLP 66

```

RESULT 10
 HPH1_RAT

```

ID  HEHL_RAT          STANDARD;          PRT;          439 AA.
AC  063244;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  HEPATOCTE NUCLEAR FACTOR 3 FORKHEAD HOMOLOG 1 (HFH-1).
GN  HFH1.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;
RX  MEDLINE=93248207; PubMed=7683413;
RA  Clevidence D.E., Overdier D.G., Tao W., Qian X., Paul L., Lai E.,
RA  Costa R.H.;
RT  "Identification of nine tissue-specific transcription factors of the
RT  hepatocyte nuclear factor 3/forkhead DNA-binding-domain family.";
RL  Proc. Natl. Acad. Sci. U.S.A. 90:3948-3952(1993).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR.
CC  -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL, L13201; AAA74561.1; -
DR  TRANSENC; T02288; -
DR  InterPro: IPR001766; -
DR  Pfam: PF00250; Fork_head. 1.
DR  PRINTS: PR00053; FORKHEAD.
DR  PROSITE: PS00657; FORK_HEAD_1; 1.
DR  PROSITE: PS00658; FORK_HEAD_2; 1.
DR  PROSITE: PS00039; FORK_HEAD_3; 1.
DR  Dna-binding; Nuclear protein; Transcription regulation.
FT  Dna BIND 99 190
FT  SEQUENCE 439 AA; 45481 MW; A37028448644C350 CRC64;
SQ

```

```

RT  four new open reading frames.";
RL  Yeast 12:599-608(1996).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL, X85811; CAA59823.1; ALT_INIT.
DR  EMBL, 271370; CAA95970.1; -
DR  SGD, S0005038; YNL094W.
RN  Hypothetical protein.
SQ  SEQUENCE 587 AA; 66134 MW; B85C525548BA34BC CRC64;
SQ

```

Query Match 62.5%; Score 40; DB 1; Length 587;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY  1 RRRPPPPYLP 10
    ||| | | |
DB  466 RRRPPPPPIP 475

```

RESULT 12
Y036_CAEEL STANDARD; PRT; 963 AA.

```

ID  Y036_CAEEL
AC  009457;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  PUTATIVE CUTICLE COLLAGEN C0965.6.
GN  C0965.6.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BRISTOL N2;
RA  Palmer S.;
RL  Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC  -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC  PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
CC  BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT (BY SIMILARITY).
CC  -1- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
CC  CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
CC  LINKS (BY SIMILARITY).
CC  -1- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
CC  COLLAGENS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL, Z46791; CAA86755.1; -
DR  WormPep: C0965.6; CE01486.
DR  InterPro: IPR000087; -
DR  InterPro: IPR002486; -
DR  Pfam: PF01484; Col_cuticle_N; 1.
DR  Pfam: PF01391; Collagen; 2.
RN  Hypothetical protein; Cuticle; Connective tissue; Repeat;
KW  Multigene family; Collagen.
FT  DOMAIN 392 423
FT  DOMAIN 441 503
FT  DOMAIN 506 567
FT  DOMAIN 506 567
FT  TRIPLE-HELICAL REGION.
FT  TRIPLE-HELICAL REGION.
FT  TRIPLE-HELICAL REGION.

```

FT DOMAIN 663 666 POLY-PRO.
 FT DOMAIN 685 688 POLY-PRO.
 SQ SEQUENCE 963 AA; 107031 MW; AFF895A75909F66E CRC64;

Query Match
 Best Local Similarity 62.5%; Score 40; DB 1; Length 963;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPYLPK 11
 1:111111
 DB 127 RRPPTMPK 136

RESULT 13
 YLL2_EBV STANDARD; PRT; 148 AA.
 AC P03199;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE HYPOTHETICAL BLF2 PROTEIN.
 CN BLF2.
 OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Gammaherpesvirinae; Lymphocryptovirus.
 NC NCB1_TaxID=10377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84270667; PubMed=6087149;
 RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
 RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Segun C.,
 RA Tuffnell P.S., Barrell B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL Nature 310:207-211(1984).
 CC -1- CAUTION: BLF2 IS KNOWN AS BLF3 IN REF.1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: V01555; CA24853.1; -
 DR PIR: A03761; Q0BE20.
 DR PIR: S33007; S33007.
 KW Hypothetical protein; Early protein.
 SQ SEQUENCE 148 AA; 16651 MW; AC4B184F0BCB391 CRC64;

Query Match
 Best Local Similarity 60.9%; Score 39; DB 1; Length 148;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPYLPK 11
 :|||||1
 DB 43 QRRPPYAKR 53

RESULT 14
 YLL2_EBV8 STANDARD; PRT; 148 AA.
 AC Q07285;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL BLF2 PROTEIN.
 CN BLF2.
 OS Epstein-Barr virus (strain AG876) (Human herpesvirus 4), and
 OS Epstein-Barr virus (strain P3HR-1) (Human herpesvirus 4).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCB1_TaxID=82830, 82829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=AG876, AND P3HR-1;
 RX MEDLINE=93331716; PubMed=8393237;
 RA Lees J.F., Arrand J.E., Pepper S.V., Stewart J.P., Mackett M.,
 RA Arrand J.R.;
 RT "The Epstein-Barr virus candidate vaccine antigen gp340/220 is highly
 RT conserved between virus types A and B.";
 RL Virology 195:578-586(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L07923; AAA02784.1; -
 DR EMBL: L07923; AAA02788.1; -
 KW Hypothetical protein; Early protein.
 SQ SEQUENCE 148 AA; 16669 MW; AC4B0953BC8A54A CRC64;

Query Match
 Best Local Similarity 60.9%; Score 39; DB 1; Length 148;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPYLPK 11
 :|||||1
 DB 43 QRRPPYAKR 53

RESULT 15
 XG_HUMAN STANDARD; PRT; 180 AA.
 AC P55808;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE XG GLYCOPROTEIN PRECURSOR (PROTEIN PBDX).
 GN XG OR PBDX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Bone marrow;
 RX MEDLINE=94332149; PubMed=8054981;
 RA Ellis N.A., Ye T.Z., Patton S., German J., Goodfellow P.N.,
 RA Weller P.;
 RT "Cloning of PBDX, an MIC2-related gene that spans the pseudoautosomal
 RT boundary on chromosome X.";
 RL Nat. Genet. 6:394-400(1994).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- PTM: O-GLYCOSYLATED (PROBABLE).
 CC -1- POLYMORPHISM: XG IS RESPONSIBLE FOR THE XG BLOOD GROUP SYSTEM.
 CC -1- SIMILARITY: TO PROTEIN MIC2/CD99.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X96421; -; NOT_ANNOTATED_CDS.
 DR MIM: 314700; -
 KW Transmembrane; Glycoprotein; Blood group antigen; Signal.

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 180 XG GLYCOPROTEIN.
FT TRANSMEM 143 163 POTENTIAL.
SQ SEQUENCE 180 AA; 19723 MW; DADA9E6859C4530 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 180;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPPYLP 11
:|||||:
Db 47 KRPYPYPO 55

RESULT 16

BCT7_SHEEP
ID BCT7_SHEEP STANDARD; PRT; 190 AA.
AC P50415;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BACNECIN 7 PRECURSOR (BAC7).
GN BAC7.5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scoochi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins.";
RT FEES Lett 376:225-228(1995).
-1- FUNCTION: EXERTS, IN VITRO, A POTENT ANTIMICROBIAL ACTIVITY.
PROBABLY DUE TO AN IMPAIRMENT OF THE FUNCTION OF THE RESPIRATORY
CHAIN AND OF ENERGY-DEPENDENT ACTIVITIES IN THE INNER MEMBRANE
OF SUSCEPTIBLE MICROORGANISMS (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; L46852; AAA85468.1; -;
DR InterPro; IPR001894; -;
DR Pfam; PF00666; Cathelicidins; 1.
DR PROSITE; PS00946; CATHELICIDINS_1; 1.
DR PROSITE; PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Repeat; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 130 BY SIMILARITY.
FT CHAIN 131 190 BACNECIN 7.
FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).
FT DISULFID 85 96 BY SIMILARITY.
FT DISULFID 107 124 BY SIMILARITY.
SQ SEQUENCE 190 AA; 21829 MW; E4AFAF1600E98371 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 190;
Best Local Similarity 72.7%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 RRRPPYLP 11
:|||||:
Db 132 RLRRPRLPR 142

RESULT 17
MTRP_HUMAN
ID MTRP_HUMAN STANDARD; PRT; 233 AA.
AC Q15012;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GOLGI 4-TRANSMEMBRANE SPANNING TRANSPORTER MTP.
GN MTRP OR KIAA0108.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Bone marrow;
RX MEDLINE=95308325; PubMed=7788527;
RA Nagase T., Miyajima K., Tanaka A., Sazuka T., Seki N., Sato S.,
RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
RT Prediction of the coding sequences of unidentified human genes. III.
RT The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by
RT analysis of cDNA clones from human cell line KG-1.*;
RT DNA Res. 2:37-43(1995).
-1- FUNCTION: MAY FUNCTION IN THE TRANSPORT OF NUCLEOSIDES AND/OR
NUCLEOSIDE DERIVATIVES BETWEEN THE CYTOSOL AND THE LUMEN OF AN
INTRACELLULAR MEMBRANE-BOUND COMPARTMENT (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MAY RESIDE IN AN
INTRACELLULAR MEMBRANE-BOUND COMPARTMENT (POTENTIAL).
-1- DOMAIN: THE C-TERMINAL DOMAIN IS NECESSARY FOR RETENTION WITHIN
INTRACELLULAR MEMBRANES (BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; D14696; BAA03522.1; -;
DR Transmembrane; Transport.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 82 102 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT DOMAIN 226 229 POLY-PRO.
SQ SEQUENCE 233 AA; 26800 MW; 9BB76152B6E1365E CRC64;

Query Match 60.9%; Score 39; DB 1; Length 233;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 RRRPPYLP 10
:|||||:
Db 224 KRPYPYLP 232

RESULT 18
MTRP_MOUSE
ID MTRP_MOUSE STANDARD; PRT; 233 AA.
AC Q60961;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GOLGI 4-TRANSMEMBRANE SPANNING TRANSPORTER (MOUSE TRANSPORTER PROTEIN)
(MTP).
GN MTRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DKA/2: TISSUE=Leukemia, and Brain;
RA MEDLINE=96199248; PubMed=8621662;
RT Hogue D.L., Ellison M.J., Young J.D., Cass C.E.;
RT "Identification of a novel membrane transporter associated with
RT intracellular membranes by phenotypic complementation in the yeast
RT Saccharomyces cerevisiae."
RL J. Biol. Chem. 271:9801-9808(1996).
CC -1- FUNCTION: MAY FUNCTION IN THE TRANSPORT OF NUCLEOSIDES AND/OR
CC NUCLEOSIDE DERIVATIVES BETWEEN THE CYTOSOL AND THE LUMEN OF AN
CC INTRACELLULAR MEMBRANE-BOUND COMPARTMENT.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MAY RESIDE IN AN
CC INTRACELLULAR MEMBRANE-BOUND COMPARTMENT (POTENTIAL).
CC -1- ALTERNATIVE PRODUCTS: A TRUNCATED ISOFORM LACKING RESIDUE 1-89
CC (MTP1) MAY BE PRODUCED BY ALTERNATIVE INITIATION. MTP1 WOULD
CC PROBABLY BE LOCALIZED IN THE PLASMA MEMBRANE.
CC -1- DOMAIN: THE C-TERMINAL DOMAIN IS NECESSARY FOR RETENTION WITHIN
CC INTRACELLULAR MEMBRANES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34259; AAB04938.1; -
DR MGD; MGI:108017; Mtrp.
KW Transmembrane; Transport.
FT TRANSMEM 29 49
FT TRANSMEM 82 102 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT DOMAIN 226 229 POLY-PRO.
SQ SEQUENCE 233 AA; 26857 MW; CAA09AC77ACE4D6E CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 1; Length 233;
Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 2 RRRPPYLP 10
: | | | | |
DB 224 KEPPPYLP 232

RESULT 19
CIT4_HUMAN STANDARD; PRT; 393 AA.
AC Q9NYG8;
FT 01-OCT-2000 (Rel. 40, Created)
FT 01-OCT-2000 (Rel. 40, Last sequence update)
FT 01-OCT-2000 (Rel. 40, Last annotation update)
DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 4 (TWIK-RELATED ARACHIDONIC ACID-
DE STIMULATED POTASSIUM CHANNEL PROTEIN) (TRAK).
GN KCNK4 OR TRAK.
CS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC PubMed=11042359;
RA Chapman C.G., Meadows H.J., Godden R.J., Campbell D.A., Duckworth M.,
RA Keisell R.E., Mudock P.R., Randall A.D., Rennie G.I., Gloger I.S.;
RT "Cloning, localisation and functional expression of a novel human,
RT cerebellum specific, two pore domain potassium channel."
RL Brain Res. Mol. Brain Res. 82:74-83(2000).
[2]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Frontal cortex;
RA Gray A.T.;
RT "Assignment of KCNK4 encoding the human potassium channel TRAK to a
RT chromosome 11."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: VOLTAGE INSENSITIVE, INSTANTANEOUS, OUTWARDLY RECTIFYING
CC POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS REVERSED AT HIGH
CC EXTERNAL K+ CONCENTRATIONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
CC CHANNELS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF248242; AAG31731.1; -
DR EMBL; AF247042; AAF64062.1; ALT_INIT.
DR InterPro: IPR001622; -
DR InterPro: IPR003280; -
DR PRINTS; PRO1333; 2PORKCHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Potassium transport;
KW Glycoprotein.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 24 POTENTIAL.
FT DOMAIN 89 113 PORE-FORMING (POTENTIAL).
FT TRANSMEM 118 138 POTENTIAL.
FT DOMAIN 140 171 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 172 192 POTENTIAL.
FT DOMAIN 197 221 PORE-FORMING (POTENTIAL).
FT TRANSMEM 234 254 POTENTIAL.
FT DOMAIN 255 393 POTENTIAL.
FT CARBOHYD 78 78 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
P -> L (IN REF. 2).
SQ SEQUENCE 393 AA; 42704 MW; 7F1B53A0A9AD57D CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 1; Length 393;
Pred. No. 45;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 RRRPPYLP 10
: | | | | |
DB 368 RRRPPPRKP 377

RESULT 20
RKG_NOTYI STANDARD; PRT; 505 AA.
AC P18516; Q04643;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RETINOIC ACID RECEPTOR GAMMA (RAR-GAMMA) (RETINOIC ACID RECEPTOR
DE DELTA) (RAR-DELTA).
GN RARG OR NR1B3.
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
CC Notophthalmus.
CC NCBI_TaxID=8316;
RN [1]
RN SEQUENCE FROM N.A. (RAR-DELTA-1 AND RAR-DELTA-2).
RP TISSUE=Tail;
RC MEDLINE=93183772; PubMed=8382939;
RA Ragsdale C.W. Jr., Gates P.B., Hill D.S., Brookes J.P.;

```

"Delta retinoic acid receptor isoform delta 1 is distinguished by its exceptional N-terminal sequence and abundance in the limb regeneration blastema.";
 Mech. Dev. 40:99-112(1993).
 [2]
 RP SEQUENCE OF 52-505 FROM N.A.
 RA Regsdale C.W. Jr., Petkovich M., Gates P.B., Chambon P., Brookes J.P.;
 RT "Identification of a novel retinoic acid receptor in regenerative tissues of the newt.";
 RL Nature 341:654-657(1989).
 RN [3]
 RP PARTIAL SEQUENCE FROM N.A.
 RA MEDLINE-89127522; PubMed-2536901;
 RX Giguere V., Ong E.S., Evans R.M., Tabin C.J.;
 RT "Spatial and temporal expression of the retinoic acid receptor in the regenerating amphibian limb.";
 RL Nature 337:566-569(1989).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, RAR-DELTA-1 (SHOWN HERE) AND RAR-DELTA-2; ARE PRODUCED BY ALTERNATIVE SPLICING. TWO FORMS DELTA 1A AND DELTA 1B ARE PRODUCED BY THE USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
 CC -1- TISSUE SPECIFICITY: RAR-DELTA-1 IS MOST ABUNDANT IN REGENERATING LIMBS, TAILS, AND THE ANTERIOR HALF OF THE LOWER JAW. RAR-DELTA-2 IS BROADLY AND UNIFORMLY DISTRIBUTED.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC N1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69944; CAA49564.1; -;
 DR EMBL: X69944; CAA49565.1; -;
 DR EMBL: X69944; CAA49562.1; ALT_INIT.
 DR EMBL: X69945; CAA49566.1; -;
 DR EMBL: X17586; CAA35603.1; -;
 DR PIR: S06124; S06124.
 DR HSSP: P10826; 1HRA.
 DR InterPro: IPR000536; -;
 DR InterPro: IPR001628; -;
 DR InterPro: IPR001723; -;
 DR InterPro: IPR003078; -;
 DR Pfam: PF00104; hormone_rec; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PRO0047; STEROIDFINGER.
 DR PRINTS: PRO0398; STRDHOMOMER.
 DR PRINTS: PRO1292; RETNOICACIDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Multigene family; Alternative splicing;
 KW Alternative initiation.
 KM CHAIN 1 505 RETINOIC ACID RECEPTOR DELTA-1A.
 FT INIT_MET 52 505 RETINOIC ACID RECEPTOR DELTA-1B.
 FT DOMAIN 52 142 MODULATING.
 FT DNA_BIND 143 208 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 143 163 C4-TYPE.
 FT ZN_FING 179 203 C4-TYPE.
 FT DOMAIN 131 137 POLY-PRO.
 FT DOMAIN 219 222 POLY-LYS.
 FT DOMAIN 419 422 POLY-ARG.

FT DOMAIN 209 254 HINGE.
 FT DOMAIN 255 474 LIGAND-BINDING.
 FT VARSPLIC 1 114 MMKFSPTASCRDGGEEPERGKGRSKRMKKEPTGSV
 FT GKEAAVAASMSSKDRICSTSTOLSHGFSPMPFAS
 FT SNMRGSPPDLTNGCAVFSFPTDLKEMASL -> MYDCH
 FT EAFMLAPHLVDYNTGACMLKRAKLSPCFGGLDPEGWOP
 FT ASLQ (IN ISOFORM RAR-DELTA-2).
 FT CONFLICT 281 281 H -> D (IN REF. 3).
 FT CONFLICT 303 303 G -> S (IN REF. 3).
 FT CONFLICT 497 497 G -> A (IN REF. 3).
 SQ SEQUENCE 505 AA; 56629 MW; 0466F549C962E262 CRC64;
 Query Match Score 38.5; DB 1; Length 505;
 Best Local Similarity 66.7%; Pred. No. 69;
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 OY 1 RRRPREVL-PR 11
 DB 420 RRRPKPYPMEPR 431
 RESULT 21
 RN RNB_HSV1M STANDARD; PRT; 161 AA.
 AC P56958;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RNA-BINDING PROTEIN (VIM21).
 OS US11.
 GN Herpes simplex virus (type 1 / strain MP).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 CC NCBI_Taxid=10307;
 RN SEQUENCE FROM N.A.
 RP Diaz J.-J.;
 RA Submitted (JUN-2000) to the SWISS-PROT data bank.
 CC -1- FUNCTION: BINDS DNA AND RNA. IS ONE OF THE MOST ABUNDANT VIRAL PROTEINS PRESENT IN CELLS LATE IN INFECTION. NOT NECESSARY FOR VIRUS VIABILITY. COMPONENT OF THE HSV-1 VIRION, WITH APPROXIMATELY 600 TO 1000 COPIES PER VIRION. BINDS SPECIFICALLY TO U134 MRNA (IN VITRO) AND SEEMS TO REGULATE THE EXPRESSION OF THAT GENE (BY SIMILARITY).
 CC -1- SUBUNIT: ASSOCIATES WITH RNA DERIVED FROM THE 60S RIBOSOMAL SUBUNITS. SEEMS TO FORM LARGE HETEROGENEOUS POLYMERS OF UP TO 200 IDENTICAL SUBUNITS IN THE CYTOPLASM (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS. FOLLOWING INFECTION, IT IS RELEASED INTO THE CELL CYTOPLASM (BY SIMILARITY).
 CC -1- DOMAIN: THE FIRST 40 AMINO ACIDS OF THE N-TERMINAL REGION IS AN EFFECTOR DOMAIN NECESSARY FOR THE TRANSACTIVATION OF HTLV-1 ENV GENE EXPRESSION, WHICH MAY INTERACT WITH CELLULAR PROTEINS (BY SIMILARITY).
 CC -1- DOMAIN: THE N-TERMINAL TETRAPEPTIDE MAY BE RESPONSIBLE FOR VIRION INCORPORATION (BY SIMILARITY).
 CC -1- DOMAIN: THE C-TERMINAL HALF, RICH IN ARG AND PRO RESIDUES, SEEMS TO BE RESPONSIBLE FOR THE RNA-BINDING ACTIVITY, AND FOR THE ASSOCIATION WITH RIBOSOMES AND THE LOCALIZATION TO THE NUCLEOLUS. THIS REGION MAY ADOPT A POLY-L-PROLINE II HELIX SECONDARY STRUCTURE (BY SIMILARITY).
 CC -1- PTM: COULD BE PHOSPHORYLATED ON SER RESIDUES BY CELLULAR KINASES (BY SIMILARITY).
 KW Late protein; DNA-binding; RNA-binding; Nuclear protein; Repeat;
 KW Phosphorylation.
 KM DOMAIN 85 156 12 X 6 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 85 90 1.
 FT REPEAT 91 96 2.
 FT REPEAT 97 102 3.
 FT REPEAT 103 108 4.
 FT REPEAT 109 114 5.
 FT REPEAT 115 120 6.
 FT REPEAT 121 126 7.

FT REPEAT 127 132 8.
 FT REPEAT 133 138 9.
 FT REPEAT 139 144 10.
 FT REPEAT 145 150 11.
 FT REPEAT 151 156 12.
 SQ SEQUENCE 161 AA; 17745 MW; 1CF68615474462FF CRC64;

Query Match 59.4%; Score 38; DB 1; Length 161;
 Best Local Similarity 70.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRRPPPLPR 11
 11111111
 Db 137 RRRPPPPRR 146

RESULT 22
 ID Y215_ADE02 STANDARD; PRT; 215 AA.
 AC P03291;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DE HYPOTHEMETICAL PROTEIN F-215.
 CS Human adenovirus type 2.
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 CC NCBI_TaxID=10515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83056843; PubMed=7142161;
 RA Gengerics T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
 RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
 RL "Nucleotide sequences from the adenovirus-2 genome.";
 RL J. Biol. Chem. 257:13475-13491(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83056844; PubMed=7142162;
 RA Alestrom P., Akusjærvi G., Pettersson M., Pettersson U.;
 RL "DNA sequence analysis of the region encoding the terminal protein
 RT and the hypothetical N-gene product of adenovirus type 2.";
 RL J. Biol. Chem. 257:13492-13498(1982).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----

CC EMBL: J01917; NOT_ANNOTATED_CDS.
 DR PIR: A03863; A03863.
 KW Hypothetical protein.

SQ SEQUENCE 215 AA; 22787 MW; 94FA33BAAB7AF137 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 215;
 Best Local Similarity 77.8%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPPL 9
 11111111
 Db 135 RRRPPPL 143

RESULT 23
 ADHA_RHIME STANDARD; PRT; 340 AA.
 ID ADHA_RHIME
 AC 031186;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE ALCOHOL DEHYDROGENASE (EC 1.1.1.1).
 GN ADHA.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymA (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RC MEDLINE=98323558; PubMed=9659380;
 RX Willis L.B., Walker G.C.;
 RA "Identification of the Rhizobium meliloti alcohol dehydrogenase gene
 RT (adha) and heterologous expression in *Alcaligenes eutrophus*.";
 RL Biochim. Biophys. Acta 1384:197-203(1998).
 CC -1- CATALYTIC ACTIVITY: ALCOHOL + NAD(+) -> ALDEHYDE OR KETONE + NADH.
 CC -1- COFACTOR: REDUCES ZINC FOR ITS ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY. HIGH, WITH OTHER BACTERIAL ADH'S.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----

DR EMBL: AF031940; AAB87463.1; -
 DR InterPro: IPR002085; -
 DR InterPro: IPR002328; -
 DR Pfam: PF00107; adh_zinc.1.
 DR PROSITE: PS00059; ADH_ZINC.1.
 KW Oxidoreductase; zinc; NAD; Plasmid.
 FT METAL 40 40 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 63 63 ZINC (CATALYTIC) (BY SIMILARITY).
 FT SEQUENCE 340 AA; 36235 MW; C1A04B5FE5C95E1 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 340;
 Best Local Similarity 62.5%; Pred. No. 55;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPPPLP 10
 11111111
 Db 54 RRRPPPLP 61

RESULT 24
 MKK1_YEAST STANDARD; PRT; 508 AA.
 ID MKK1_YEAST
 AC P32490;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROTEIN KINASE MKK1/SSP32 (EC 2.7.1.-).
 GN MKK1 OR SSP32 OR YOR231W OR O5095.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93233668; PubMed=8386320;
 RA Irie K., Takase M., Lee K.S., Levin D.E., Araki H., Matsumoto K.,
 RA Oshima Y.;
 RT "MKK1 and MKK2, which encode Saccharomyces cerevisiae
 RT mitogen-activated protein kinase-kinase homologs, function in the
 RT pathway mediated by protein kinase C.";
 RL Mol. Cell. Biol. 13:3076-3083(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;

RA MEDLINE-97127829; PubMed-8972580;
 RA Boyer J., Michaux G., Fairhead C., Gaillon L., Dujon B.;
 RA "Sequence and analysis of a 26.9 kb fragment from chromosome XV of
 RA the yeast *Saccharomyces cerevisiae*.";

RL Yeast 12.1575-1586(1996).

CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
 CC TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
 CC AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
 CC INVOLVE THE KINASE PKC1 THAT MAY ACT ON THE BCK1 KINASE THAT THEN
 CC PHOSPHORYLATES MKK1 AND MKK2 WHICH THEMSELVES PHOSPHORYLATE THE
 CC MKP1 KINASE.

CC -1- FUNCTION: KINASE INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT
 CC INVOLVE THE PROTEIN KINASES ENCODED BY PKC1, BCK1, AND MKP1.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: D13001; BA02364.1; -;
 DR EMBL: Z75139; CA959451.1; -;
 DR PIR: S30772; S30772.
 DR PIR: A48069; A48069.
 DR HSSP: P11362; IEG1.
 DR SGD: S0005757; MKK1.
 DR InterPro: IPR000719; -;
 DR InterPro: IPR002290; -;
 DR Pfam: PF00069; Pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase: Tyrosine-protein kinase;
 KW ATP-binding; Phosphorylation.
 FT DOMAIN 221 488 PROTEIN KINASE.
 FT NP_BIND 227 235 ATP (BY SIMILARITY).
 FT BINDING 250 250 ATP (BY SIMILARITY).
 FT ACT_SITE 349 349 BY SIMILARITY.
 FT MOD_RES 377 377 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 381 381 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 508 AA; 56720 MW; 4FE62CDD185CDAA2 CRC64;

QY 1 RRRRPPLP 10
 Db 81 KRRPAPSLP 90

Query Match 59.4%; Score 38; DB 1; Length 508;
 Best Local Similarity 60.0%; Pred. No. 82;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 25
 CP84.ARATH STANDARD; PRT; 520 AA.
 ID CP84.ARATH
 AC 042600;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CYTOCHROME P450 84A1 (PERULATE-5-HYDROXYLASE) (EC 1.14.-.-) (F5H).
 DE CYP84A1 OR FAH1 OR AT4G36220 OR F23E13.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;

RX MEDLINE-96293440; PubMed-8692910;
 RA Meyer K., Cusumano J.C., Somerville C.R., Chapple C.C.S.;
 RA "Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new
 RA family of cytochrome P450-dependent monooxygenases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6869-6874(1996).

CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=CV. COLUMBIA;
 CC MEDLINE-20083488; PubMed-10617198;
 CC Meyer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 CC Pohl T., Duesternoeft A., Stiekema W., Ertan K.-D., Terry N.,
 CC Harris B., Ansoyge W., Brandt P., Obermayer B., Mache R., Mueller M.,
 CC Weichselgartner M., de Simone V., Obermayer B., Mache R., Mueller M.,
 CC Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
 CC Reichert B., Portetelle D., Perez-Alonso M., Botry M., Bancroft I.,
 CC Vos P., Hobeisel J., Zimmermann W., Medler H., Ridley P.,
 CC Langham S.-A., McCullagh B., Billham L., Robben J.,
 CC Van der Schueren J., Grymoprez B., Chuang J.-J., Vandebussche F.,
 CC Breken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 CC Weitzenecker T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
 CC Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,
 CC Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 CC Benneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 CC de Keyser A., Buysnaert C., Gielen J., Villarroel R., De Clercq R.,
 CC Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 CC Clark L., Doggett J., Hall S., Kay M., Leonard N., McIay K., Mayes R.,
 CC Pettett A., Rajandream M.-A., Lyne M., Benes V., Reckmann S.,
 CC Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 CC Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Aner S.,
 CC Gabel C., Fuchs M., Farman B., Granderath K., Dauner D., Heitzl A.,
 CC Neumann S., Argitlou A., Vitale D., Liguori R., Piravandi E.,
 CC Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 CC Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 CC Chedford F., Cooke R., Berger C., Montfort A., Casacuberta E.,
 CC Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 CC Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 CC Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 CC Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 CC Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Habermann K.,
 CC Parnell L., Dedbia N., Gnoj L., Schurz K., Huang E., Spiegel L.,
 CC Sekhon M., Murray J., Sheet P., Cordes M., Abu-Thaiden J.,
 CC Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 CC Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 CC Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 CC Kramer J., Fulton J., Mardis E., Dante M., Pepin K., Hiller L.,
 CC Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 CC Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 CC Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 CC Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 CC Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 CC Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
 CC Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis*
 RT *thaliana*.";
 RL Nature 402:769-777(1999).

CC [3]
 CC SEQUENCE FROM N.A.
 CC STRAIN=CV. LANDSBERG ERRECTA;
 CC MEDLINE-99097044; PubMed-9880351;
 CC Ruegger M., Meyer K., Cusumano J.C., Chapple C.C.;
 CC "The regulation of ferulate-5-hydroxylase expression in *Arabidopsis* in
 CC the context of sinapate ester biosynthesis.";
 CC Plant Physiol. 119:101-110(1999).

CC -1- PATHWAY: GENERAL PHENYLPROPANOID PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
ER EMBL; U38416: AAC9389.1; -
ER EMBL; AL022141: CAA18128.1; -
ER EMBL; AF068574: AAD11580.1; -
DR EMBL; AL161589: CAB80293.1; -
DR InterPro: IPR001128; -
DR Pfam: PF00067: p450.1;
DR PRINTS: PR00385; P450.1;
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
RW Oxidoreductase; Monooxygenase; Membrane; Heme.
FT BINDING 458 458 HEME (BY SIMILARITY).
SQ SEQUENCE 520 AA; 58720 MW; E812779AF5BF01BC CRC64;

Query Match 59.4%; Score 38; DB 1; Length 520;
Best Local Similarity 77.8%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 RRRPPYLP 10
    |||||
    34 RRRPPYPP 42

RESULT 26
PPSA_AERPE STANDARD; PRT; 845 AA.
AC Q9TEC5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PHOSPHOENOLPYRUVATE SYNTHASE (EC 2.7.9.2) (PYRUVATE, WATER DIKINASE)
DE (PEP SYNTHASE).
CN PPSA OR APE0650.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RT SEQUENCE FROM N.A.
RA STRAIN=K1.
RA MEDLINE=99310339; PubMed=10382966;
RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankel A., Kosugi H.,
RA Hosoya A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kusuda N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
*Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.;
RL DNA Res. 6:83-101(1999)
CC -1- CATALYTIC ACTIVITY: ATP + PYRUVATE + H(2)O = AMP +
CC PHOSPHOENOLPYRUVATE + PHOSPHATE.
CC -1- PATHWAY: ESSENTIAL STEP IN GLUCOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -1- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000060; BAA79621.1; -
CC InterPro: IPR00121; -
CC InterPro: IPR002192; -
CC Pfam: PF00391; PEP-utilizers; 1.
CC PROSITE: PS00326; PPD_N term; 1.
CC PROSITE: PS00370; PEP_ENZYMES_PHOS_SITE; 1.
CC PROSITE: PS00742; PEP_ENZYMES_2; 1.
RW Transferase; Kinase; ATP-binding; Phosphorylation.
FT MOD_RES 463 463 PHOSPHORYLATION (BY SIMILARITY).

```

```

SQ SEQUENCE 845 AA; 94799 MW; 55C3CD78E2BCB9A5 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 845;
Best Local Similarity 70.0%; Pred. No. 1,4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 2 RRRPPYLP 11
    |||||
    DB 115 RTRPPYLR 124

RESULT 27
IRSL_HCMVA STANDARD; PRT; 846 AA.
ID IRSL_HCMVA
AC P09715;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN HQR1.
GN IRSL.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RT SEQUENCE FROM N.A.
RA MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RA "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. Mol. Biol. 192:177-208(1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=90269039; PubMed=2161319;
RX Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RX Horsnell T., Hutchison C.A. III, Kourazides T., Martignetti J.A.,
RA Preddie E., Satchell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17403; CAA35311.1; -
CC EMBL; X04650; CAA28312.1; -
DR PIR: C26078; OOBEC3.
DR PIR: S09914; S09914.
DR HSSP: P41249; IAKP.
KW Hypothetical protein.
FT CARBOHYD 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 91048 MW; 622B6CFE211674BD CRC64;

Query Match 59.4%; Score 38; DB 1; Length 846;
Best Local Similarity 77.8%; Pred. No. 1,4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 3 RRRPPYLP 11
    |||||
    DB 824 RRRPPAPR 832

RESULT 28

```

RRG2_HUMAN STANDARD; PRT; 443 AA.
 AC P22932;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-2 (RAR-GAMMA-2).
 GN RARG OR NR1B3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE OF 1-63 FROM N.A.
 RA MEDLINE=90207264; PubMed=2157210;
 RA Kastner P., Krust A., Mendelsohn C., Garnier J.M., Zelent A.,
 Leroy P., Staub A., Chambon P.;
 RT "Murine isoforms of retinoic acid receptor gamma with specific
 RT patterns of expression."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2700-2704(1990).
 [2]
 RP SEQUENCE OF 51-443 FROM N.A.
 RA MEDLINE=89315787; PubMed=2546152;
 RA Krust A., Kastner P., Petkovich M., Zelent A., Chambon P.;
 RT "A third human retinoic acid receptor, hRAR-gamma."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5310-5314(1989).
 [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 167-412.
 RA Renaud J.-P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H.,
 Moras D.;
 RT "Crystal structure of the RAR-gamma ligand-binding domain bound to
 RT all-trans retinoic acid."
 RL Nature 378:681-689(1995).
 [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 167-412.
 RA MEDLINE=96162555; PubMed=9501913;
 RA Klaholz B.P., Renaud J.-P., Mitschler A., Zusi C., Chambon P.,
 Gronemeyer H., Moras D.;
 RT "Conformational adaptation of agonists to the human nuclear receptor
 RT RAR gamma."
 RL Nat. Struct. Biol. 5:199-202(1998).
 [1]
 RP FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 RP HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 RP A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 RP CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 [2]
 RP SUBCELLULAR LOCATION: NUCLEAR.
 [3]
 RP ALTERNATIVE PRODUCTS: RAR-GAMMA-1 AND -2 ARE PRODUCED BY
 RP ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY DIFFER IN THEIR
 RP N-TERMINAL SECTION.
 [4]
 RP DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 RP A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 [5]
 RP SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 [6]
 RP NBI SUBFAMILY.
 [7]
 RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RP the European Bioinformatics Institute. There are no restrictions on its
 RP use by non-profit institutions as long as the content is in no way
 RP modified and this statement is not removed. Usage by and for commercial
 RP entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RP or send an email to license@sib-sib.ch).
 [8]
 RP EMBL; M24857; NOT ANNOTATED_CDS.
 RP EMBL; M32074; AAA60253.1; -
 RP PDB; 1LBD; 02-MAR-99.
 RP PDB; 4LBD; 02-MAR-99.
 RP TRANSFAC; T00720; -
 RP MIM; 180190; -
 RP InterPro; IPR000536; -
 RP InterPro; IPR001628; -
 RP InterPro; IPR001723; -
 RP InterPro; IPR003078; -

DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STRDINGER.
 DR PRINTS; PR00398; STRDINGER.
 DR PRINTS; PR01292; RETINOICACIDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Multigene family; Alternative splicing; 3D-structure.
 FT DOMAIN 1 78 MODULATING.
 FT DNA_BIND 79 144 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 79 99 C4-TYPE.
 FT ZN_FING 115 139 C4-TYPE.
 FT DOMAIN 145 190 HINGE.
 FT DOMAIN 191 410 LIGAND-BINDING.
 SQ SEQUENCE 443 AA; 49307 MW; 4D709194F511E86 CRC64;
 Query Match 58.6%; Score 37.5; DB 1; Length 443;
 Best Local Similarity 66.7%; Pred. No. 85;
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 1 RRRPPPYL-PR 11
 DB 356 RRRPSQPYMPFR 367
 RESULT 29
 RRG1_HUMAN STANDARD; PRT; 454 AA.
 AC P13631;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-1 (RAR-GAMMA-1).
 GN RARG OR NR1B3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89315787; PubMed=2546152;
 RA Krust A., Kastner P., Petkovich M., Zelent A., Chambon P.;
 RT "A third human retinoic acid receptor, hRAR-gamma."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5310-5314(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91042616; PubMed=2172793;
 RA Ishikawa T., Umesono K., Mangelsdorf D.J., Aburatani H., Staenger B.Z.,
 Shibasaki Y., Imawari M., Evans R.M., Takaku F.;
 RT "A functional retinoic acid receptor encoded by the gene on human
 RT chromosome 12."
 RL Mol. Endocrinol. 4:837-844(1990).
 [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91187677; PubMed=1849262;
 RA Lehmann J.M., Hoffmann B., Pfahl M.;
 RT "Genomic organization of the retinoic acid receptor gamma gene."
 RL Nucleic Acids Res. 19:573-578(1991).
 [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 178-423.
 RA Renaud J.-P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H.,
 Moras D.;
 RT "Crystal structure of the RAR-gamma ligand-binding domain bound to
 RT all-trans retinoic acid."
 RL Nature 378:681-689(1995).
 [5]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 178-423.
 RA MEDLINE=96162555; PubMed=9501913;
 RA Klaholz B.P., Renaud J.-P., Mitschler A., Zusi C., Chambon P.,
 Gronemeyer H., Moras D.;
 RT "Conformational adaptation of agonists to the human nuclear receptor

RT RAR gamma.";
 CC Nat. Struct. Biol. 5:199-202(1998).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 CC HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 CC A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CC CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: RAR-GAMMA-1 AND -2 ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY DIFFER IN THEIR
 CC N-TERMINAL SECTION.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M24857; AAA52692.1; -
 CC EMBL: M38258; AAA60254.1; -
 CC EMBL: M57707; AAA63254.1; -
 CC PIR: A33903; A33903.
 CC PIR: A35573; A35573.
 CC PIR: S26848; S26848.
 CC PDB: 3LBD; 02-MAR-99.
 CC PDB: 4LBD; 02-MAR-99.
 CC TRANSFAC: T00720; -
 CC TRANSFAC: T01330; -
 CC MIM: 180190; -
 CC InterPro: IPR000536; -
 CC InterPro: IPR001628; -
 CC InterPro: IPR001723; -
 CC InterPro: IPR003078; -
 CC Pfam: PF00104; hormone_rec; 1.
 CC Pfam: PF00105; zf-C4; 1.
 CC PRINTS: PR00047; STROIDFINGER.
 CC PRINTS: PR00398; STROHORMONER.
 CC PRINTS: PR01292; RETINOICACIDR.
 CC PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
 CC KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger; Multigene family; Alternative splicing; 3D-structure.
 CC FT DOMAIN 1 89 MODULATING.
 CC FT DNA_BIND 90 155 NUCLEAR RECEPTOR-TYPE.
 CC FT ZN_FING 90 110 C4-TYPE.
 CC FT ZN_FING 126 150 C4-TYPE.
 CC FT DOMAIN 156 201 HINGE.
 CC FT DOMAIN 202 421 LIGAND-BINDING.
 CC SEQUENCE 454 AA; 50341 MW; 1EE27B22772D4AFD CRC64;
 QY 1 RRRPRPYL-PR 11
 DB 367 RRRPSQPMFPR 378
 Query Match 58.6%; Score 37.5; DB 1; Length 454;
 Best Local Similarity 66.7%; Pred. No. 87;
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OS Streptomyces coelicolor.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.
 CC NCBI_TaxID=1902;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE-95189113; PubMed-7883195;
 RA Urahe H., Ogawara H.;
 RT "Cloning, sequencing and expression of serine/threonine kinase-
 RL encoding genes from Streptomyces coelicolor A3(2).";
 RN gene 153:99-104(1995).
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,
 RA Rajadream M.A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- PTM: AUTOPHOSPHORYLATED MAINLY AT THR AND SLIGHTLY AT SER.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D86821; BA013168.1; -
 CC EMBL: AL163672; CAB87324.1; -
 CC HSSP: 063450; 1A06.
 CC InterPro: IPR000719; -
 CC InterPro: IPR001898; -
 CC InterPro: IPR002290; -
 CC Pfam: PF00069; pkinase; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC KW transferase; Serine/threonine-protein kinase; ATP-binding;
 CC Phosphorylation.
 CC FT DOMAIN 8 276 PROTEIN KINASE.
 CC FT NP_BIND 14 22 ATP (BY SIMILARITY).
 CC FT BINDING 48 48 ATP (BY SIMILARITY).
 CC FT ACT_SITE 142 142 BY SIMILARITY.
 CC FT DOMAIN 301 481 GLN/PRO-RICH.
 CC SEQUENCE 543 AA; 58181 MW; 0E1965520FA0C200 CRC64;
 QY 1 RRRP-----RPPYLP 11
 DB 402 RRRRQVSRPQPR 417
 Query Match 58.6%; Score 37.5; DB 1; Length 543;
 Best Local Similarity 56.2%; Pred. No. 1e+02; 2; Indels 5; Gaps 1;
 Matches 9; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```

RC STRAIN-168;
RA MEDLINE-97124191; PubMed-8969504;
RX Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,
RA Sanders J., Emerson P.T., Hatwood C.R.;
RT "the dnaB-phaA (256 degrees-240 degrees) region of the Bacillus
subtilis chromosome containing genes responsible for stress
responses," the utilization of plant cell walls and primary
metabolism."
RL Microbiology 142:3067-3078(1996).
CC -1- CATALYTIC ACTIVITY: L-ARABINOSE = L-RIBULOSE.
CC -1- PATHWAY: FIRST STEP OF L-ARABINOSE CATABOLISM AFTER IT IS
CC TRANSPORTED INTO THE CELL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z75208; CA99587.1; -.
CC EMBL: Z99118; CAB14840.1; -.
CC Subtilist; BG11904; araa.
CC Arabiose catabolism; isomerase.
CC SEQUENCE 498 AA; 56540 MW; D6D50EB90514494C CRC64;
SQ

Query Match 57.8%; Score 37; DB 1; Length 498;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
DB 90 KRREPPY 97

RESULT 32
FKBX_MOUSE STANDARD; PRT; 581 AA.
ID FKX_MOUSE
AC 061576;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 65 KDA FK506-BINDING PROTEIN PRECURSOR (EC 5.2.1.8) (FKBP65) (FKBPRP)
DE (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE) (PIPIASE) (ROTAMASE)
DE (IMMUNOPHILIN FKBP65).
DE FKBP6 OR FKBP1-RS OR FKBP-RS OR FKBP-RP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JB6: TISSUE-Epidermis;
RX MEDLINE-94117013; PubMed-7507077;
RA Simek S.L., Kozak C.A., Winterstein D., Hegamyer G., Colburn N.H.;
RT "Sequence and localization of a novel FK506-binding protein to mouse
RT chromosome 11."
RL Genomics 18:407-409(1993).
RN [2]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE-96094328; PubMed-7493967;
RA Coss M.C., Winterstein D., Sowder R.C. II, Simek S.L.;
RT "Molecular cloning, DNA sequence analysis, and biochemical
RT characterization of a novel 65-kDa FK506-binding protein (FKBP65).";
RL J. Biol. Chem. 270:29336-29341(1995).
CC -1- FUNCTION: PIPIASES ACCELERATE THE FOLDING OF PROTEINS DURING
CC PROTEIN SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -1- ENZYME REGULATION: INHIBITED BY BOTH FK506 AND RAPAMYCIN, BUT NOT
CC BY CYCLOSPORINE A.

```

```

CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE LUNG, SPLEEN, HEART, BRAIN
CC AND TESTIS.
CC -1- PTM: GLYCOSYLATED AND PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PIPIASE FAMILY. CONTAINS 4
CC FKBP-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L07063; AAC37678.1; -.
CC MGD: MGI:104769; Fkbp6.
CC InterPro: IPR000886; -.
CC InterPro: IPR001179; -.
CC InterPro: IPR002048; -.
CC Pfam: PF00254; FKBP_4.
CC Pfam: PF00036; efhand; 2.
CC PROSITE: PS00453; FKBP_PIPIASE_1; FALSE_NEG.
CC PROSITE: PS00454; FKBP_PIPIASE_2; 1.
CC PROSITE: PS00059; FKBP_PIPIASE_3; 4.
CC PROSITE: PS00014; ER_TARGET; 1.
CC PROSITE: PS00018; EF_HAND; 1.
CC Isomerase; Rotamase; Repeat; Glycoprotein; Signal; Phosphorylation;
CC Endoplasmic reticulum; Calcium-binding.
CC SIGNAL 1 33
CC CHAIN 34 581
CC DOMAIN 54 146
CC DOMAIN 179 258
CC DOMAIN 291 370
CC DOMAIN 389 482
CC CA_BIND 509 520
CC CA_BIND 554 565
CC CARBOHYD 69 69
CC CARBOHYD 181 181
CC CARBOHYD 293 293
CC CARBOHYD 309 309
CC CARBOHYD 351 351
CC CARBOHYD 392 392
CC CARBOHYD 406 406
CC SITE 578 581
CC SEQUENCE 581 AA; 64669 MW; 1B51B3032089F555 CRC64;
SQ

Query Match 57.8%; Score 37; DB 1; Length 581;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 RPPYLPR 11
DB 162 RPPYCR 168

RESULT 33
SG2_BOVIN STANDARD; PRT; 613 AA.
ID SG2_BOVIN
AC P20616;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SECRETORANIN II PRECURSOR (SGII) (CHROMOGHRANIN C).
GN SCG2 OR CHG2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA MEDLINE-90264409; PubMed-2345170;
 RA Fischer-Colbrie R., Gutierrez J., Hsu C.M., Iacangelo A., Eiden L.E.;
 RT "Sequence analysis, tissue distribution and regulation by cell
 RT depolarization, and second messengers of bovine secretogranin II
 RT (chromogranin C) mRNA.";
 RL J. Biol. Chem. 265:9208-9213(1990).
 CC -1- FUNCTION: SECRETORANIN II IS A NEUROENDOCRINE SECRETORY GRANULE
 CC PROTEIN, WHICH IS THE PRECURSOR FOR BIOLOGICALLY ACTIVE PEPTIDES.
 CC -1- SUBCELLULAR LOCATION: NEUROENDOCRINE AND ENDOCRINE SECRETORY
 CC GRANULES.
 CC -1- MISCELLANEOUS: BINDS CALCIUM WITH A LOW-AFFINITY.
 CC -1- SIMILARITY: BELONGS TO THE CHROMOGRANIN / SECRETORANIN PROTEIN
 CC FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL: J05468; AAA30760.1; -
 DR PIR: A35296; A35296.
 DR InterPro: IPR001990; -
 DR Pfam: PF01271; Granin; 1.
 DR PROSITE: PS00422; GRANIN_1; 1.
 DR Sulfatation: Cleavage on pair of basic residues; Calcium-binding;
 MW Sulfatation.
 FT SIGNAL. 1 27 POTENTIAL.
 FT PROPEP 28 30 SECRETORANIN II.
 FT CHAIN 31 613 UNKNOWN ACTIVITY PEPTIDE (PROBABLE).
 FT PEPTIDE 181 213 SULFATATION (BY SIMILARITY).
 FT MOD.RES 150 150
 FT MOD.RES 150 150
 SQ SEQUENCE 613 AA; 70356 MW; 5DC079F59D83516 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 613;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PRPPYLP 10
 DB 444 PKRPPKPR 450

RESULT 34
 VGLB_HSV11 STANDARD; PRT; 904 AA.
 AC P10211;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GLYCOPROTEIN B PRECURSOR.
 GN GB OR U27.
 OS Herpes simplex virus (type 1 / strain 17).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 CC NCBI_TaxId=10299;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-88274327; PubMed-2839594;
 RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
 RA McNeib D., Perry L.J., Scott J.E., Taylor P.;
 RT "The complete DNA sequence of the long unique region in the genome of
 RT herpes simplex virus type 1.";
 RL J. Gen. Virol. 69:1531-1574(1988).
 CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
 CC GB, GC, GG, GD, GI, AND GE.
 CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
 CC REQUIRED FOR VIRAL GROWTH.

CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL: D10879; BAA01673.1; -
 DR EMBL: X14112; CAA32320.1; -
 DR PIR: I30084; VGBEW7.
 DR InterPro: IPR000234; -
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR Glycoprotein; Transmembrane; Signal.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL. 1 30
 FT CHAIN 31 904 GLYCOPROTEIN B.
 FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 731 746 POTENTIAL.
 FT TRANSMEM 752 771 POTENTIAL.
 FT TRANSMEM 775 795 POTENTIAL.
 FT DOMAIN 796 904 POTENTIAL.
 FT CARBOHYD 87 87 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 904 AA; 100292 MW; 2C14E8B1284C1F3A CRC64;

Query Match 57.8%; Score 37; DB 1; Length 904;
 Best Local Similarity 60.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRRPPYLP 11
 DB 73 RRRPPKPR 82

RESULT 35
 PCK2_SCHPO STANDARD; PRT; 1016 AA.
 ID PCK2_SCHPO
 AC P36583;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROTEIN KINASE C-LIKE 2 (EC 2.7.1.-).
 GN PCK2 OR STS6 OR PKC1 OR SPBCL2D12.04C.
 OS Schizosaccharomyces pombe (Fission Yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI_TaxId=4896;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-93259141; PubMed-8491190;
 RA Toda T., Shimanuki M., Yanagida M.;
 RT "Two novel protein kinase C-related genes of fission yeast are
 RT essential for cell viability and implicated in cell shape control.";
 RL EMBO J. 12:1987-1995(1993).
 CC [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-93216684; PubMed-8463273;
 RA Mazzei G.J., Schmid E.M., Knowles J.K., Payton M.A., Maundrell K.G.;
 RT "A Ca(2+)-independent protein kinase C from fission yeast.";
 RL J. Biol. Chem. 268:7401-7406(1993).
 CC [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
 RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN THE CONTROL OF THE CELL SHAPE. TARGET
 CC OF THE INHIBITOR STAUROSPORINE.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PRC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D14338; BA003268.1; -
 CC EMBL: L07637; AAA35323.1; -
 CC EMBL: AL035085; CAA22678.1; -
 CC PIR: S35363; S35363.
 CC PIR: A46079; A46079.
 CC HSSP: P05132; 2CFK.
 CC InterPro: IPR000719; -
 CC InterPro: IPR000961; -
 CC InterPro: IPR002219; -
 CC Pfam: PF00130; DAG_PE-bind; 2.
 CC Pfam: PF00433; pkinase; 1.
 CC ProSite: PS00479; DAG_PE_BIND_DOM_1; 2.
 CC ProSite: PS00081; DAG_PE_BIND_DOM_2; 2.
 CC ProSite: PS00107; PROTEIN_KINASE_ATP; 1.
 CC ProSite: PS00108; PROTEIN_KINASE_DOM; 1.
 CC ProSite: PS00108; PROTEIN_KINASE_ST; 1.
 CC Transfester; Serine/threonine-protein kinase; ATP-binding; Zinc;
 CC Phorbol-ester binding; Duplication.
 FT DOMAIN 406 453 PHORBOL-ESTER AND DAG BINDING (BY
 FT DOMAIN 474 523 PHORBOL-ESTER AND DAG BINDING (BY
 FT DOMAIN 554 560 POLY-HIS.
 FT DOMAIN 683 942 PROTEIN KINASE.
 FT NP_BIND 689 697 ATP (BY SIMILARITY).
 FT BINDING 712 712 ATP (BY SIMILARITY).
 FT ACT_SITE 808 808 E -> V (IN REF. 1).
 FT CONFLICT 152 152
 SQ SEQUENCE 1016 AA; 116004 MW; 060CDC4F718A0069 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 1016;
 Best Local Similarity 55.6%; Pred. No. 2.3e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPYLP 10
 DB 953 KRTGPPYIP 961

RESULT 36
 GTPA_HUMAN STANDARD; PRT; 1047 AA.
 AC P20936;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GTPASE-ACTIVATING PROTEIN (GAP) (RAS P21 PROTEIN ACTIVATOR) (P120GAP)
 DE (RASGAP).
 GN RASAI OR RASA.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;

RX MEDLINE=89072759; PubMed=3201259;
 RA Trahey M., Wong G., Halenbeck R., Rubinfeld B., Martin G.A.,
 RA Ladner M., Long C.M., Crosier W.J., Watt K., Kothe K.,
 RA McCormick F.;
 RT "Molecular cloning of two types of GAP complementary DNA from human
 RT placenta.";
 RL Science 242:1697-1700(1988).
 RN [2]
 RP STRUCTURE BY NMR OF 275-350.
 RX MEDLINE=94185632; PubMed=8137811;
 RA Yang Y.S., Garbay C., Duschesne M., Cornille F., Jullian N.,
 RA Fromage N., Tocque B., Rogues B.P.;
 RT "Solution structure of GAP SH3 domain by 1H NMR and spatial
 RT arrangement of essential Ras signaling-involved sequence.";
 RL EMBD J. 13:1270-1279(1994).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 714-1047.
 RX MEDLINE=97113417; PubMed=8955277;
 RA Scheffzek K., Lautwein A., Kabsch W., Reza Ahmadian M.,
 RA Wittinghofer A.;
 RT "Crystal structure of the GTPase-activating domain of human p120GAP
 RT and implications for the interaction with Ras.";
 RL Nature 384:591-596(1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 714-1047 IN COMPLEX WITH RAS.
 RX MEDLINE=97364817; PubMed=9219684;
 RA Scheffzek K., Ahmadian M.R., Kabsch W., Wiesmuller L., Lautwein A.,
 RA Schmitz F., Wittinghofer A.;
 RT "The Ras-RasGAP complex: structural basis for GTPase activation and
 RT its loss in oncogenic Ras mutants.";
 RL Science 277:333-338(1997).
 RN [5]
 RP VARIANTS LEU-398; GLY-400 AND VAL-401.
 RX MEDLINE=94100977; PubMed=8275088;
 RA Friedman E., Gejman P.V., Martin G.A., McCormick F.;
 RT "Nonsense mutations in the C-terminal SH2 region of the GTPase
 RT activating protein (GAP) gene in human tumours.";
 RL Nat. Genet. 5:242-247(1993).
 CC -1- FUNCTION: GAP STIMULATES THE GTPASE OF NORMAL BUT NOT ONCOGENIC
 CC RAS P21.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. ARE PROBABLY PRODUCED BY ALTERNATIVE SPLICING
 CC AND CAUSE BASAL CELL CARCINOMAS.
 CC -1- DISEASE: MUTATIONS IN THE SH2 DOMAIN OF RASA SEEM TO BE ONCOGENIC
 CC -1- SIMILARITY: TO OTHER RAS GTPASE-ACTIVATING PROTEINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M23379; AAA52517.1; -
 CC EMBL: M23612; AAA35865.1; -
 CC PIR: A40121; A40121.
 CC PDB: IWER; 3I-DEC-97.
 CC PDB: IWO1; 15-JUL-98.
 CC MIM: 139150; -
 CC InterPro: IPR000008; -
 CC InterPro: IPR000980; -
 CC InterPro: IPR001452; -
 CC InterPro: IPR001849; -
 CC InterPro: IPR001936; -
 CC Pfam: PF00168; C2; 1.
 CC Pfam: PF00169; PH; 1.
 CC Pfam: PF00616; RasGAP; 1.

Query Match 57.8%; Score 37; DB 1; Length 1342;
 Best Local Similarity 63.6%; Pred. No. 3.1e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPPYLPR 11
 111 11: 11
 Db 1203 RRRHSPHPPR 1213

RESULT 38
 P300_HUMAN STANDARD; PRT; 2414 AA.
 AC 009472;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE EIA-ASSOCIATED PROTEIN P300.
 GN EP300 OR P300.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95011587; PubMed=7523245;
 RA Eckner R., Even M.E., Newsome D., Gerdes M., Decaprio J.A.,
 Lawrence J.B., Livingston D.M.;
 RA "Molecular cloning and functional analysis of the adenovirus EIA-
 associated 300-kD protein (p300) reveals a protein with properties of
 a transcriptional adaptor.";
 RL Genes Dev. 8:869-884(1994).
 CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL ADAPTOR REQUIRED FOR THE
 ACTIVITY OF CERTAIN COMPLEX TRANSCRIPTIONAL REGULATORY ELEMENTS.
 CC MAY HAVE A FUNCTION IN CELL CYCLE REGULATION. BINDS TO AND MAY BE
 INVOLVED IN THE TRANSFORMING CAPACITY OF THE ADENOVIRUS EIA
 PROTEIN.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U01877; AAA18639.1; -;
 DR MIM; 602700; -;
 DR InterPro; IPR000197; -;
 DR InterPro; IPR000433; -;
 DR InterPro; IPR001487; -;
 DR Pfam; PF00569; 22; 1.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF02135; zf-TAZ; 2.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR PROSITE; PS00663; BROMODOMAIN_1; 1.
 DR PROSITE; PS50014; BROMODOMAIN_2; 1.
 KW Transcription regulation; Nuclear protein; Bromodomain; Cell cycle.
 FT DOMAIN 11 17 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 797 800 POLY-SER.
 FT DOMAIN 1067 1139 BROMODOMAIN.
 FT DOMAIN 1519 1526 POLY-GLU.
 FT DOMAIN 1572 1818 BINDING REGION FOR EIA ADENOVIRUS.
 FT DOMAIN 2066 2069 POLY-GLN.
 FT DOMAIN 2190 2195 POLY-GLN.
 SQ SEQUENCE 2414 AA; 264143 MW; 6BFF909E4B9D693 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 2414;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 6 PPYLR 11
 11111
 Db 1884 PPYLR 1889

RESULT 39
 RRG2_XENLA STANDARD; PRT; 442 AA.
 AC P28699; P79879; P79878;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-2 (RAR-GAMMA-2).
 GN RARG OR NR1B3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91115103; PubMed=1846602;
 RA Ellinger-Ziegelbauer H., Dreyer C.;
 RT "A retinoic acid receptor expressed in the early development of
 Xenopus laevis.";
 RL Genes Dev. 5:94-104(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96191149; PubMed=8641047;
 RA Crawford M.J., Liverage R.A., Varmuza S.L.;
 RT "Two isoforms of Xenopus retinoic acid receptor gamma 2 (b) exhibit
 differential expression and sensitivity to retinoic acid during
 embryogenesis.";
 RL Dev. Gene. 17:291-302(1995).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 HAS DEEP EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CELL FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; RAR-GAMMA-1 AND RAR-GAMMA-2;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THEIR N-
 TERMINAL SECTION.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X59396; CAA42039.1; -;
 DR EMBL; S82175; AAB47116.1; -;
 DR EMBL; S82173; AAB47115.1; -;
 DR PIR; A38592; A38592.
 DR HSSP; P10826; 1HRA.
 DR TRANSFAC; T01358; -;
 DR InterPro; IPR000536; -;
 DR InterPro; IPR001628; -;
 DR InterPro; IPR001723; -;
 DR InterPro; IPR003078; -;
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-CA; 1.
 DR PRINTS; PR00047; STROIDINGER.
 DR PRINTS; PR00398; STRDHOMONER.
 DR PRINTS; PR01292; RETNOICACIDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.

Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 zinc-finger; Multigene family; Alternative splicing.

FT DNABIND 1 75 MODULATING.
 FT ZN_FING 76 141 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 76 96 C4-TYPE.
 FT DOMAIN 112 136 C4-TYPE.
 FT DOMAIN 142 187 HINGE.
 FT DOMAIN 188 408 LIGAND-BINDING.
 FT DOMAIN 150 155 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT CONFLICT 7 7 A -> S (IN ABA47115).
 FT CONFLICT 208 208 Q -> R (IN REF. 2).
 FT CONFLICT 347 347 G -> A (IN REF. 2).
 FT SEQUENCE 442 AA; 49889 MW; 27D65665D48BC22C CRC64;

Query Match 57.0%; Score 36.5; DB 1; Length 442;
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 1 RRRPPPYL-PR 11
 ||||| : : : : :
 Db 354 RRRPKPYMFPR 365

RESULT 40
 RGL_XENLA STANDARD: PRT; 476 AA.

AC P51127;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-1 (RAR-GAMMA-1).
 GN RARG OR NR183
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 ON NCBI_TaxID=8355;
 RX MEDLINE-92196110; PubMed-131217;
 RA Blumberg B., Mangelsdorf D.J., Dyck J.A., Bittner D.A., Evans R.M.,
 de Robertis E.M.;
 RT "Multiple retinoid-responsive receptors in a single cell: families of
 retinoid 'X' receptors and retinoic acid receptors in the xenopus
 egg".
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2321-2325(1992).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; RAR-GAMMA-1 AND RAR-GAMMA-2; ARE
 PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THEIR N-
 TERMINAL SECTION.
 CC -1- DEVELOPMENTAL STAGE: IT IS SYNTHESIZED DURING OOGENESIS AND
 PERSISTS DURING EARLY CLEAVAGE. IT ACCUMULATES AT GASTRULATION
 (STAGE 10), PEAKS IN QUANTITY DURING NEURULATION (STAGE 17), THEN
 DROPS TO A LOW LEVEL AFTER STAGE 26.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L11444; -; NOT_ANNOTATED_CDS.

HSP; P10826; 1HRA.
 DR InterPro; IPR000536; -;
 DR InterPro; IPR001628; -;
 DR InterPro; IPR001723; -;
 DR InterPro; IPR003078; -;
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STRDIPFINGER.
 DR PRINTS; PR00398; STRDHOMER.
 DR PRINTS; PR01292; RETINOICACID.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Multigene family; Alternative splicing.

FT DNABIND 1 109 MODULATING.
 FT ZN_FING 110 175 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 110 130 C4-TYPE.
 FT ZN_FING 146 170 C4-TYPE.
 FT DOMAIN 176 221 HINGE.
 FT DOMAIN 222 442 LIGAND-BINDING.
 FT DOMAIN 184 189 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT SEQUENCE 476 AA; 53033 MW; 2AB772BB0448F7FF CRC64;

Query Match 57.0%; Score 36.5; DB 1; Length 476;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 1 RRRPPPYL-PR 11
 ||||| : : : : :
 Db 388 RRRPKPYMFPR 399

RESULT 41
 SSS2_SCYCA STANDARD: PRT; 80 AA.

AC P11020;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE SPERMATID-SPECIFIC PROTEIN S2 (BASIC NUCLEAR PROTEIN S2).
 OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Galeidae; Carcharhiniformes;
 OC Scyllorhinidae; Scyllorhinus.
 ON NCBI_TaxID=7830;
 RX MEDLINE-89170733; PubMed-2924768;
 RA Chauviere M., Martingue A., Briand G., Sautiere P.,
 RT "Nuclear basic protein transition during sperm differentiation.
 Primary structure of the spermatid-specific protein S2 from the
 dog-fish Scyllorhinus canicula".
 RL Eur. J. Biochem. 180:329-335(1989).
 CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES
 ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
 REPLACED BY PROTEINS IN LATE SPERMATIDS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- MISCELLANEOUS: N-TERMINAL HALF IS HIGHLY BASIC, WHILE C-TERMINAL
 PART IS ACID.
 CC -1- SIMILARITY: WITH PROTEIN S1 (70% STRUCTURAL SIMILARITY).
 CC PIR; S03560; S03560.
 DR Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Nuclear protein.
 KW SEQUENCE 80 AA; 9733 MW; 593E18EF3452409 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 80;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

OY 1 RRRPPPYL-PR 10
 ||||| : : : : :
 Db 26 RRRPKPYMFPR 39

```

RESULT 42
ID GT22_ARATH STANDARD; PRT: 221 AA.
AC G92V03;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE GLUTATHIONE S-TRANSFERASE ZETA-CLASS 2 (EC 2.5.1.18).
GN AT2G02390 OR T1FF16.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Beil C.R., Ketchum K.A., Lee J.U., Rongning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-G.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ZETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC005312; AAC78521.1;
CC DR Mendel: 33869; Arach:1313;33869.
CC DR InterPro: IPR000521;
CC DR Pfam: PF00043; GST; 1.
CC KM Transferase. 17
CC FT ACT_SITE 17
CC FT SEQUENCE 221 AA; 24887 MW; 106BDC8EF3E745BF CRC64;
SQ

```

```

RN
RP [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90156494; PubMed=2154589;
RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
RT repeats, is conserved in several limited passage isolates but not in
RT strain 17syn+."
RL J. Virol. 64:1014-1020(1990).
CC -1- FUNCTION: NOT KNOWN. HAS A PROFOUND EFFECT ON THE GROWTH OF THE
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33700; AAA45791.1;
CC Repeat: Late protein.
CC FT DOMAIN 155
CC FT SEQUENCE 245 AA; 25981 MW; 55AD450AB1B8CA11 CRC64;
SQ

```

Query Match 56.2%; Score 36; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 RRRP 7
   |||||
Db 12 RRRP 17

```

```

RESULT 44
ID ICP3_HSV11 STANDARD; PRT: 248 AA.
AC P36313;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
GN ICP34.5.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=92341080; PubMed=1321882;
RA Dolan A., McKie E., McLean A.R., McGeoch D.J.;
RT "Status of the ICP34.5 gene in herpes simplex virus type 1 strain
RT 17."
RL J. Gen. Virol. 73:971-973(1992).
CC -1- FUNCTION: NOT KNOWN. HAS A PROFOUND EFFECT ON THE GROWTH OF THE
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X14112; -; NOT_ANNOTATED_CDS.
 DR PIR: J01682; J01682.

KM Repeat.
 KM DOMAIN 161 175 5 X 3 AA TANDEM REPEATS OF A-T-P.
 SQ SEQUENCE 248 AA: 26185 MW; AD334AF0C9CABA65 CRC64;

Query Match 56.2%; Score 36; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 |||||
 Db 11 RRRPP 16

RESULT 45
 ICP3_HSV1D STANDARD; PRT; 252 AA.
 ID ICP3_HSV1D

AC P37318;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
 GN ICP34.5.
 OS Herpes simplex virus (type 1 / strain CVC-2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=37106;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90156494; PubMed=2154589;

RA Chou J., Roizman B.;
 RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
 RT repeats, is conserved in several limited-passage isolates but not in
 RT strain 17syn+.";
 RL J. Virol. 64:1014-1020(1990).

CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33701; AAA45792.1; -;
 KM Repeat. Late protein.
 FT DOMAIN 162 179 6 X 3 AA TANDEM REPEATS OF A-T-P.
 SQ SEQUENCE 252 AA: 26601 MW; A6E6D049FEAA79A CRC64;

Query Match 56.2%; Score 36; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 |||||
 Db 12 RRRPP 17

Search completed: September 24, 2001, 10:09:06
 Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 24, 2001, 10:07:41 : Search time 47.47 Seconds
(without alignments)
30.658 Million cell updates/sec

Title: US-09-276-868-4
Perfect score: 64
Sequence: 1 RRRPRPYLR 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL_16:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP:invertebrate:*
7: SP:mammal:*
8: SP:mhc:*
9: SP:organelle:*
10: SP:phage:*
11: SP:plant:*
12: SP:rodent:*
13: SP:unclassified:*
14: SP:vertebrate:*
15: SP:virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	71.9	184	5 Q23291	Q23291 caenorhabd
2	45.5	71.1	602	14 Q66852	Q66852 fowl adenov
3	45	70.3	200	2 Q9RKS4	Q9RKS4 streptomyc
4	45	70.3	278	13 Q42232	Q42232 columba liv
5	45	70.3	304	6 Q9N1Q3	Q9N1Q3 canis fami
6	45	70.3	354	2 Q05915	Q05915 mycobacteri
7	45	70.3	361	2 Q9XCG4	Q9XCG4 mycobacteri
8	45	70.3	372	6 Q9N1Q2	Q9N1Q2 canis fami
9	44	68.8	575	4 Q00235	Q00235 homo sapien
10	44	68.8	575	6 P79282	P79282 sus scrofa
11	44	68.8	575	6 Q9N0W2	Q9N0W2 bos taurus
12	44	68.8	619	5 Q9WTS2	Q9WTS2 mus musculu
13	44	68.8	675	10 Q9LPR4	Q9LPR4 drosophila
14	44	68.8	675	10 Q9LPR4	Q9LPR4 arachidopsi
15	43	67.2	590	4 Q00456	Q00456 homo sapien
16	43	67.2	652	14 Q10436	Q10436 m dna polym
17	42.5	66.4	177	10 Q9FSP4	Q9FSP4 oryza sativ
18	42	65.6	589	5 Q18756	Q18756 caenorhabd
19	42	65.6	740	5 Q21426	Q21426 caenorhabd

20	42	65.6	915	5 Q9W267	Q9W267 drosophila
21	41	64.1	89	2 Q9RE11	Q9RE11 lactobacill
22	41	64.1	104	10 Q9S1Z9	Q9S1Z9 arachidopsi
23	41	64.1	133	10 Q9SM70	Q9SM70 oryza sativ
24	41	64.1	136	10 Q9FW62	Q9FW62 oryza sativ
25	41	64.1	168	10 Q9SM77	Q9SM77 oryza sativ
26	41	64.1	289	5 Q9W5C6	Q9W5C6 drosophila
27	41	64.1	300	10 Q9LW27	Q9LW27 oryza sativ
28	41	64.1	665	10 Q23664	Q23664 arachidopsi
29	41	64.1	2328	5 Q9VW47	Q9VW47 drosophila
30	41	64.1	2531	5 Q9GPH4	Q9GPH4 drosophila
31	40	62.5	31	2 Q30708	Q30708 bacillus su
32	40	62.5	129	10 Q24156	Q24156 nicotiana t
33	40	62.5	262	2 Q06242	Q06242 mycobacteri
34	40	62.5	287	2 Q9F345	Q9F345 streptomyc
35	40	62.5	301	10 Q41848	Q41848 zea mays (m
36	40	62.5	324	10 Q81299	Q81299 arachidopsi
37	40	62.5	400	11 Q70220	Q70220 mus musculu
38	40	62.5	400	11 Q9JLN7	Q9JLN7 rattus ratt
39	40	62.5	400	11 Q9JLN7	Q9JLN7 mus musculu
40	40	62.5	402	4 Q9NS06	Q9NS06 homo sapien
41	40	62.5	596	5 Q9VNA6	Q9VNA6 drosophila
42	40	62.5	600	2 Q9HVQ5	Q9HVQ5 pseudomonas
43	40	62.5	693	13 Q91889	Q91889 xenopus lae
44	40	62.5	701	13 Q91890	Q91890 xenopus lae
45	39	60.9	134	1 Q9Y9Q5	Q9Y9Q5 aeropyrum p

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	184 AA.
Q23291	Q23291			
AC	Q23291			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-NOV-1998 (TREMBLrel. 08, Last annotation update)			
DE	COSMID ZC404.			
GN	ZC404.1.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RX	MEDLINE=94150718; PubMed=7906398;			
RA	Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,			
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,			
RA	Craighton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,			
RA	Gardner A., Green P., Hawkins T., Hillier L., Jiler M., Johnston L.,			
RA	Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,			
RA	Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,			
RA	Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,			
RA	Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,			
RA	Therly-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,			
RA	Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.			
RT	elegans."			
RL	Nature 368:32-38(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RA	Bentley D., Le T., T.,			
RL	Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2.			
RA	Waterston R.,			
RL	Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: U55563; AAA97967.1; -			
SO	SEQUENCE 184 AA; 21649 MW; 6D64D09EA91308F9 CRC64;			

```
Query Match
Best Local Similarity 71.9%; Score 46; DB 5; Length 184;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPPYLP 11
    ||| |||
DB 26 RRRPPYLP 34

RESULT 2
Q66852 PRELIMINARY; PRT; 602 AA.
AC Q66852;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE PRE-TERMINAL PROTEIN.
OS fowl adenovirus.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviaadenovirus.
OX NCBI_TaxID=31540;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE 10;
RA McCoy R.J., Sheppard M.;
RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sheppard M., Werner W., McCoy R., Johnson M.A.;
RL Arch. Virol. 0:0-0(1997);
DR EMBL; U26220; AAB02174.1; -.
DR EMBL; AF007577; AAB88665.1; -.
DR InterPro; IPR003391; -.
DR Pfam; PF02459; Adeno_terminal; 1.
SQ SEQUENCE 602 AA; 70392 MW; EB48277B4390894F CRC64;

Query Match
Best Local Similarity 71.1%; Score 45.5; DB 14; Length 602;
Matches 10; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY 1 RRRPP-----YLP 10
    ||| |||
DB 338 RRRPPSPQPEEYLP 354

RESULT 3
Q9RK54 PRELIMINARY; PRT; 200 AA.
AC Q9RK54;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
DE HYPOTHETICAL 22.1 KDA PROTEIN.
GN SCF12.02C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
```

```
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL117669; CAB56128.1; -.
KW Hypothetical protein.
SQ SEQUENCE 200 AA; 22076 MW; 0DCBBEC585803B5 CRC64;

Query Match
Best Local Similarity 70.3%; Score 45; DB 2; Length 200;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPYLP 11
    ||| |||
DB 118 RRRPPYLP 128

RESULT 4
Q42232 PRELIMINARY; PRT; 278 AA.
AC Q42232;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE Y-BOX BINDING PROTEIN A (FRAGMENT).
RN YB-A.
OS Columba livia (Domestic pigeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX NCBI_TaxID=8932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pratt S.L., Horseman N.D.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF022769; AAB80761.1; -.
DR HSSP; P41016; IC90.
DR InterPro; IPR002059; -.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PRO0050; COLDSHOCK.
DR SMART; SM00357; CSP; 1.
KW DNA-binding; Transcription regulation.
FT NON_TER
SQ SEQUENCE 278 AA; 29614 MW; B51D26D012DF9410 CRC64;

Query Match
Best Local Similarity 70.3%; Score 45; DB 13; Length 278;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRPPYLP 10
    ||| |||
DB 242 RRRPPYLP 251

RESULT 5
Q9N1Q3 PRELIMINARY; PRT; 304 AA.
AC Q9N1Q3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE Y-BOX PROTEIN ZONAB-A.
GN ZONAB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
```


RX MEDLINE-20253078; PubMed-10790369;
RA Balda M.S., Matter K.;
RT "the tight junction protein ZO-1 and an interacting transcription
factor regulate ErbB-2 expression.";
RL EMO J. 19:2024-2033(2000).
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF171061; AAF72335.1; -.
DR InterPro; IPR002059; -.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLD_SHOCK.
DR PROSITE; PS00352; COLD_SHOCK; 1.
DR SMART; SM00357; CSP; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 304 AA; 31939 MW; C49E737720F5202D CRC64;

Query Match 70.3%; Score 45; DB 6; Length 304;
Best Local Similarity 80.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 RRRPPPLP 10
|||||
Db 268 RRRPPNAP 277

RESULT 6
ID 005915 PRELIMINARY; PRT; 354 AA.
AC 005915;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
ET 01-JUL-1997 (TREMBLrel. 04, last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
EE HYPOTHEICAL 39.1 KDA PROTEIN.
GN RV0923C OR MTCY21C12.17C.
CS Mycobacterium tuberculosis.
CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxID=1773;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-96342230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeller K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsey T., Jagels K., Krogh A., Molean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
Taylor K., Whitehead S., Barrell B.G.; -.
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; Z95210; CAB08505.1; -.
DR Tuberculist; RV0923C; -.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39059 MW; 2A0AD48C506A5233 CRC64;

Query Match 70.3%; Score 45; DB 2; Length 354;
Best Local Similarity 88.9%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 3 RPPPLP 11
|||||
Db 111 RPPPLP 119

RESULT 7
ID 09XCG4 PRELIMINARY; PRT; 361 AA.
AC 09XCG4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, last annotation update)
DE HYPOTHEICAL 40.2 KDA PROTEIN.
OS Mycobacterium avium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1764;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2151;
RA Eckstein T.M., Lambert M.L., Brennan P.J., Bellisle J.T., Iamline J.M.;
RT "Identification of a gene cluster involved in glycopeptidolipid
biosynthesis and of a gene cluster encoding daunorubicin resistance in
two strains of Mycobacterium avium serovar 2.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF143772; AAD44199.1; -.
KW Hypothetical protein.
SQ SEQUENCE 361 AA; 40208 MW; AD01DBE825C1C9EA CRC64;

Query Match 70.3%; Score 45; DB 2; Length 361;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 RRRPPPLP 10
|||||
Db 32 RRRPPNAP 41

RESULT 8
ID 09N102 PRELIMINARY; PRT; 372 AA.
AC 09N102;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE Y-BOX PROTEIN ZONAB-B.
GN ZONAB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20253078; PubMed-10790369;
RA Balda M.S., Matter K.;
RT "The tight junction protein ZO-1 and an interacting transcription
factor regulate ErbB-2 expression.";
RL EMO J. 19:2024-2033(2000).
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF171062; AAF72336.1; -.
DR InterPro; IPR002059; -.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLD_SHOCK.
DR Prodom; PD00621; -; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
DR SMART; SM00357; CSP; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 372 AA; 39879 MW; 33E579077CF2B969 CRC64;

Query Match 70.3%; Score 45; DB 6; Length 372;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 RRRPPPLP 10
|||||
Db 336 RRRPPNAP 345

RESULT 9
ID 000235 PRELIMINARY; PRT; 575 AA.

AC 000235;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE N-ACETYL-BETA-D-GLUCOSAMINIDE (ALPHA1-6FUCOSYLTRANSFERASE).
 GN ALPHA1-6FUCT OR FUT8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97279058; PubMed=9133635;
 RA Yamaguchi S., Uozumi N., Ihara Y., Miyoshi E., Yamaguchi N.,
 RA Taniguchi N.;
 RT "Purification and cDNA cloning of GDP-L-Fuc:N-acetyl-beta-D-
 RT glucosaminide:alpha1-6 fucosyltransferase (alpha1-6 FucT) from human
 RT gastric cancer MKN45 cells.";
 RU J. Biochem. 121:626-632(1997).
 RN [2]
 RP SEQUENCE OF 278-419 FROM N.A.
 RA Yamaguchi Y.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=WHOLE EMBRYO;
 RA Cailliau A., Balanzino L., Candelier J.J., Oriol R., Mollicone R.;
 RL "Differential splice variants of human FUT8 embryonic cDNA.";
 RT Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; D89289; BAA19764.1; -;
 DR EMBL; AF038281; AAB92372.1; -;
 DR EMBL; AF038280; AAB92372.1; JOINED.
 DR EMBL; Y17976; CAA76985.1; -;
 DR EMBL; Y17976; CAA76985.1; -;
 DR EMBL; Y17977; CAA76986.1; -;
 DR EMBL; Y17978; CAA76987.1; -;
 DR InterPro: IPR001452; -;
 DR Pfam: PF00018; SH3; 1.
 DR SMART: SM00326; SH3; 1.
 KW Transferase: Glycosyltransferase.
 KM SEQUENCE 575 AA; 66515 MW; 5AE24A93881E18D0 CRC64;
 SQ
 Query Match 68.8%; Score 44; DB 4; Length 575;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 PRPPYLP 10
 Db 299 PRPPYLP 305
 RESULT 10
 P79282 PRELIMINARY; PRT; 575 AA.
 AC P79282;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE ALPHA-(1,6)-FUCOSYLTRANSFERASE (EC 2.4.1.68) (GDP-FUCOSE--GLYCOPROTEIN
 DE FUCOSYLTRANSFERASE)
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=BRIN;
 RX MEDLINE=97066976; PubMed=8910378;
 RA Uozumi N., Yamaguchi S., Miyoshi E., Ihara Y., Sakuma T., Gao C.-X.,
 RA Teshima Y., Fujii S., Shiba T., Taniguchi N.;
 RT "Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl-
 RT beta-D-glucosaminide alpha1->6fucosyltransferase.";

RL J. Biol. Chem. 271:27810-27817(1996).
 CC -1- FUNCTION: TRANSFERS FUCOSE FROM GDP-FUC TO N-LINKED TYPE COMPLEX
 CC GLYCOPETIDES. THE PH OPTIMUM IS 7.
 CC -1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + N-(N-ACETYL-BETA-D-
 CC GLUCOSAMINYL-1,2-ALPHA-D-MANNOSYL-1,3-(R(1))-ALPHA-1,6)-BETA-D-
 CC MANNOYL-BETA-N-ACETYL-1,4-D-GLUCOSAMINYL-1,4-N-ACETYL-D-
 CC GLUCOSAMINYL)ASPARAGINE = GDP + N-(N-ACETYL-BETA-D-GLUCOSAMINYL-
 CC 1,2-ALPHA-D-MANNOSYL-1,3-(R(1))-ALPHA-1,6)-BETA-D-MANNOSYL-1,4-BETA
 CC -N-ACETYL-D-GLUCOSAMINYL-1,4-(ALPHA-L-FUCOSYL-1,6)-N-ACETYL-D-
 CC GLUCOSAMINYL)ASPARAGINE.
 CC -1- PATHWAY: GLYCOSYLATION.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN.
 DR EMBL; D86723; BAA13157.1; -;
 DR InterPro: IPR001452; -;
 DR Pfam: PF00018; SH3; 1.
 DR PROSITE, PS50002; SH3; 1.
 DR SMART; SM00326; SH3; 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;
 KW Golgi stack.
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 10 25 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 26 575 LUMENAL, CATALYTIC (POTENTIAL).
 SQ SEQUENCE 575 AA; 66229 MW; 0F19D0BC2018F7B CRC64;
 Query Match 68.8%; Score 44; DB 6; Length 575;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 PRPPYLP 10
 Db 299 PRPPYLP 305
 RESULT 11
 Q9N0W2 PRELIMINARY; PRT; 575 AA.
 ID Q9N0W2;
 AC Q9N0W2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE 6-ALPHA-L-FUCOSYLTRANSFERASE.
 GN FUT8.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Javaud C., Dupuy F., Mafiah A., Michalski J.C., Oriol R., Petit J.M.,
 RA Julien R.;
 RT "Ancestral exonic organization of FUT8, the gene encoding the alpha6-
 RT fucosyltransferase reveals successive peptide domains which suggest a
 RT particular 3D core-structure for the alpha6-fucosyltransferase
 RT family.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF247186; AAF65460.1; -;
 DR InterPro: IPR001452; -;
 DR PROSITE, PS50002; SH3; 1.
 KW Transferase; Glycosyltransferase.
 KM SEQUENCE 575 AA; 66492 MW; FE04C66B8A5B540 CRC64;
 SQ
 Query Match 68.8%; Score 44; DB 6; Length 575;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 PRPPYLP 10
 Db 299 PRPPYLP 305

Db 299 PRPYLP 305

RESULT 12

Q9WTS2 PRELIMINARY; PRT; 575 AA.

AC Q9WTS2: 01-NOV-1999 (TREMblrel. 12, Created)

DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)

DE ALPHA-1,6-FUCOSYLTRANSFERASE.

GN FUT8.

CS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=CEREBRUM;

RA Hayashi H., Yoneda A., Imanura T.:

RT "Mouse alpha-1,6-fucosyltransferase.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

EMBL: AB025198; BAA76392.1; -

DR MGD: MGI:1858901; Fut8.

DR InterPro: IPR001452; -

DR Pfam: PF00018; SH3; 1.

DR SMART: SM00326; SH3; 1.

DR Transferase; Glycosyltransferase.

SQ SEQUENCE 575 AA; 6655 MW; 7BE2ED3146EC45F CRC64;

Query Match Best Local Similarity 68.8%; Score 44; DB 11; Length 575; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 4 PRPYLP 10

Db 299 PRPYLP 305

RESULT 13

Q9YVVS PRELIMINARY; PRT; 619 AA.

AC Q9YVVS: 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)

DE CG2448 PROTEIN.

GN CG2448.

OS Drosophila melanogaster (Fruit fly).

CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA MEDLINE-20196006; PubMed-10731132.

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer V., Chao M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Adali J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouch J., Brokstein P., Brotler P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de la Piedad B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M., Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong P., Correia J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C., Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein J., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Glibos R.A., Myers E.W., Rubin G.M., Venter J.C.:

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003487; AAF48079.1; -

DR Flybase: FBgn0030327; CG2448.

DR InterPro: IPR001452; -

DR Pfam: PF00018; SH3; 1.

DR SMART: SM00326; SH3; 1.

SQ SEQUENCE 619 AA; 70205 MW; 691BFD9B5C657DE CRC64;

Query Match Best Local Similarity 68.8%; Score 44; DB 5; Length 619; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 4 PRPYLP 10

Db 345 PRPYLP 351

RESULT 14

G9LPR4 PRELIMINARY; PRT; 675 AA.

AC G9LPR4: 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)

DE F15H18.3.

OS Arabidopsis thaliana (Mouse-ear cress).

CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mkharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Federpiel N.A., Theologis A., Ecker J.R.:

RT "Genomic sequence for Arabidopsis thaliana BAC F15H18 from chromosome I.";

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Ecker J.R.:

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Ecker J.R.:

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

[4]
 RN SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bel B., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A.,
 RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
 RA Ecker J.;
 RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaverl A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC013354; AAF26002.1; -
 DR InterPro: IPR000891; -
 DR Pfam: PF00682; HMG1-like; 1.
 DR PROSITE: PS00815; ALPM_HOMOCIT_SYNTH_1; 1.
 SQ SEQUENCE 675 AA; 73526 MW; 9357AA728898A6B CRC64;

Query Match 68.8%; Score 44; DB 10; Length 675;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RRRPPPLP 10
 |||||
 Db 72 RRRPPPLP 80

RESULT 15
 ID 000456 PRELIMINARY; PRT: 590 AA.
 AC 000456;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE TTF-I INTERACTING PEPTIDE 21 (FRAGMENT).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Jansa P., Grunmt I.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF000561; AAB58414.1; -
 DR InterPro: IPR000210; -
 DR Pfam: PF00651; BTB; 1.
 DR PROSITE: PS50097; BTB; 1.
 DR SMART: SM00225; BTB; 1.
 FT NON_TER 1
 SQ SEQUENCE 590 AA; 62116 MW; C8DB72DFB670067E CRC64;

Query Match 67.2%; Score 43; DB 4; Length 590;
 Best Local Similarity 80.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RRRPPPLP 11
 |||||
 Db 571 RRRPPPLP 580

RESULT 16

010436
 ID 010436 PRELIMINARY; PRT: 652 AA.
 AC 010436;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE DNA POLYMERASE (E2B DNA POL) AND PTP (E2B PTP) GENES, PARTIAL CDS, AND
 DE 52,55K (L1 52,55K), PIIIA (L1 PIIIA), III (L2 III), PVII (L2 PVII),
 DE PUTATIVE 30 KDA PROTEIN (L2 COMP. ORF), V (L2 V) AND PX (L2 PX) GENES,
 DE COMPLETE CDS (FRAGMENT).
 GN E2B PTP.
 GN E2B PTP.
 OS Mouse adenovirus type 1 (MAV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 RX NCBI_TaxID=10530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Weisner J.D., Hirsch G.N., Larue E.A., Fulcher R.A., Spindler K.R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U95843; AAB53751.1; -
 DR InterPro: IPR003391; -
 DR Pfam: PF02459; Adeno_terminal; 1.
 FT NON_TER 1
 SQ SEQUENCE 652 AA; 74663 MW; 1400CF2D2E266D0B CRC64;

Query Match 67.2%; Score 43; DB 14; Length 652;
 Best Local Similarity 80.0%; Pred. No. 57;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RRRPPPLP 10
 |||||
 Db 387 RRRPPPLP 396

RESULT 17
 ID 09FSP4 PRELIMINARY; PRT: 177 AA.
 AC 09FSP4;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 19.2 KDA PROTEIN.
 GN H0711G06.6.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 RX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Mu J., Zhou B., Chen Z.H., Li Y., Zhu J.J.,
 RA Tang Y.S., Zhao Q., Liu Y.L., Huang Y.C., Yu Z., Fan D.L., Chen L.,
 RA Weng Q.J., Zhang L., Lu Y.O., Yu S.L., Zhu J., Liu X.H., Hu X.,
 RA Lei H.Y., Zhang Y.J., Wang R., Li C., Lu Y., Chen X.C., Zhang Y.,
 RA Hu H., Jia P.X., Li T., Qian Y.M., Ying K., Hong G.F.;
 RT "Oryza sativa indica (Guanglu14) genomic DNA, chromosome 4, BAC
 clone: H0711G06 (+H013C06).";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL442115; CAC09500.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 177 AA; 19223 MW; 7F4BACF1A04B41E6 CRC64;

Query Match 66.4%; Score 42.5; DB 10; Length 177;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 QY 1 RRRPP-PLP 11
 |||||
 Db 54 RRRPP-PLP 65

RESULT 18

Q18756
ID Q18756 PRELIMINARY: PRT: 589 AA.
AC Q18756:
CT 01-NOV-1996 (TREMBlrel. 01, Created)
ET 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, last annotation update)
DE PROLINE AND GLYCINE-RICH.
GN C50F7.2.
CS Caenorhabditis elegans.
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
CC Rhabditidae; Pelodierinae; Caenorhabditis.
CX NCBI_TaxID=6239;
FN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
FA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
FA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
FA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
FA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
FA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
FA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
FA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
FA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
FA Talery-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
FA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
FT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
FT elegans.";
RL Nature 368:32-38(1994).
RN [12]
RP SEQUENCE FROM N.A.
RA Johnson D., Steillyes L.;
RA Submitted (DDBJ-1995) to the EMBL/GenBank/DBJ databases.
RN [13]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U41557; AAA83307.1; -;
DR InterPro: IPR000087; -;
SQ SEQUENCE 589 AA; 55491 MW; 038508B521A5EB9 CRC64;

Query Match 65.6%; Score 42; DB 5; Length 589;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPPRYLPR 11
Db 105 RRRPPRYLPR 114

RESULT 19
ID Q21426 PRELIMINARY: PRT: 740 AA.
AC Q21426: Q21627;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE R01H10.8 PROTEIN.
DE R01H10.8.
CS Caenorhabditis elegans.
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
CC Rhabditidae; Pelodierinae; Caenorhabditis.
CX NCBI_TaxID=6239;
FN [1]
RP SEQUENCE FROM N.A.
RA Mortimore B.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z31590; CAA83466.1; -;
DR InterPro: IPR001478; -;
DR InterPro: IPR001660; -;
DR InterPro: IPR001849; -;
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00595; PD2; 1.

DR SMART: SM00454; SAM; 1.
SQ SEQUENCE 740 AA; 83904 MW; FFC0B6CAB78210BF CRC64;

Query Match 65.6%; Score 42; DB 5; Length 740;
Best Local Similarity 63.6%; Pred. No. 91;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPRYLPR 11
Db 368 RRRPPRYLPR 378

RESULT 20
ID Q9W267 PRELIMINARY: PRT: 915 AA.
AC Q9W267:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE CG5819. PROTEIN.
CN CG5819.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burks K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Clewley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Fouts K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Sprung R., Smith T.,
RA Spier E., Spidling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venier E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
FT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003457; AAF46828.1; -;
DR HSSP: P23945; 1XUN.
DR FlyBase: FBgn0034717; CG5819.
DR InterPro: IPR000483; -.

DR InterPro; IPR001611; --
 DR Pfam; PF00560; LRR; 9.
 DR Pfam; PF01463; LRCT; 1.
 DR PRINTS; PRO0019; LEURICRPT.
 DR SMART; SMO0082; LRCT; 1.
 SQ SEQUENCE 915 AA; 103360 MW; 4B9A2B71A767C4F6 CRC64;

Query Match 65.6%; Score 42; DB 5; Length 915;
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPPPLP 10
 : : : : :
 Db 784 KRKPPLP 793

RESULT 21
 ID Q9RE11 PRELIMINARY; PRT; 89 AA.
 AC Q9RE11;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 10.6 KDA PROTEIN.
 OS Lactobacillus casei.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Lactobacillus.
 OC NCBI_Taxid=1582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 393;
 RA Dossoumet V., Deutscher J.;
 RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; Y18948; CAB65149.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 89 AA; 10573 MW; E7EED1D5E7BFD668 CRC64;

Query Match 64.1%; Score 41; DB 2; Length 89;
 Best Local Similarity 72.7%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRPPPLP 11
 : : : : :
 Db 56 RRRPPPLP 66

RESULT 22
 ID Q9S1Z9 PRELIMINARY; PRT; 104 AA.
 AC Q9S1Z9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE AR2G22000 PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 Buell C.R., Ketchum K.A., Lee J.D., Ronning C.M., Koo H., Moffat K.S.,
 Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
 Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
 Salzberg S.L., Fraser C.M., Venter J.C.;

RT "Sequence and analysis of chromosome II of Arabidopsis thaliana."
 RL Nature 402:761-768(1999).
 DR EMBL; AC007019; AAD20417.1; --
 SQ SEQUENCE 104 AA; 11633 MW; B00A462F6C7ED91E CRC64;

Query Match 64.1%; Score 41; DB 10; Length 104;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPP 7
 : : : : :
 Db 87 RRRPP 93

RESULT 23
 ID Q9SM70 PRELIMINARY; PRT; 133 AA.
 AC Q9SM70;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE ZHB0015.1
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OC NCBI_Taxid=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. INDICA;
 RA Hong G., Zhou B.;
 RT "Oryza sativa genomic DNA, chromosome 4, clone: t17804."
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL117265; CAB55423.1; --
 SQ SEQUENCE 133 AA; 14869 MW; 7DBBD3D19EF91A3 CRC64;

Query Match 64.1%; Score 41; DB 10; Length 133;
 Best Local Similarity 77.8%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RRRPP 11
 : : : : :
 Db 117 RRRPP 125

RESULT 24
 ID Q9FW62 PRELIMINARY; PRT; 136 AA.
 AC Q9FW62;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 14.4 KDA PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OC NCBI_Taxid=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanden S.E.,
 Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblyum T.V.,
 Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 10 BAC OSJNB0094K03 genomic sequence."
 RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC069145; AAG16854.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 136 AA; 14435 MW; D3ACG33BEB362C91 CRC64;

```

Query Match          64.1%; Score 41; DB 10; Length 136;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRRPRPP 7
    |||||
Db 99 RRRPRPP 105

RESULT 25
O9SM77 PRELIMINARY; PRT; 168 AA.
AC O3SM77;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE ZH00008.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC Oryza.
NC NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, INDICA;
RA Hong S., Zhou B.;
RT "Oryza sativa genomic DNA, chromosome 4, clone: t17804.";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
SQ SEQUENCE 168 AA; 18668 MW; B9A19B483840F1A3 CRC64;

Query Match          64.1%; Score 41; DB 10; Length 168;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRRPRPP 7
    |||||
Db 56 RRRPRPP 62

RESULT 26
O9W5C6 PRELIMINARY; PRT; 289 AA.
AC O9W5C6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG14632.
OS CG14632.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006, PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulten G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modyar G., Morris J., Moshnell A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
DR EMBL: AE003419; AAF45561.1;
DR FlyBase: FBgn0029538; CG14632.
SQ SEQUENCE 289 AA; 31446 MW; CE232C2D6868A1B0 CRC64;

```

```

Query Match          64.1%; Score 41; DB 5; Length 289;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RRRPRPP 10
    |||||
Db 177 RRRPRPP 186

RESULT 27
O9LWZ7 PRELIMINARY; PRT; 300 AA.
ID O9LWZ7;
AC O9LWZ7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE SIMILAR TO ORYZA SATIVA ROOT-SPECIFIC RCC3 mRNA.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
OC Oryza.
NC NCBI_TaxID=4530;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(ga3) genomic DNA, chromosome 6, PAC
clone:p0644B06.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001129; BAA90617.1;
DR InterPro: IPR000528;
DR InterPro: IPR003612;
DR Pfam: PF00279; LTR; 1.
DR SMART: SM00499; AAI; 1.
SQ SEQUENCE 300 AA; 30134 MW; B3A7C2E42CA3425A CRC64;

```

```

Query Match          64.1%; Score 41; DB 10; Length 300;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRRPRPP 7
    |||||

```

DB 178 RRRPRPP 184

RESULT 28

023664 PRELIMINARY; PRT: 665 AA.

AC 023664;

DT 01-JAN-1998 (TREMBlrel. 05, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE AUXIN RESPONSE FACTOR 1.

GN AREF1 OR F23H11.7.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RX Ulasov T., Hagen G., Guilfoyle T.J.;

RT "AREF1, a transcription factor that binds to auxin response elements.";

RL Science 276:1865-1868(1997).

RU [2]

RN SEQUENCE FROM N.A.

RC STRAIN=CV. COLOMBIA;

RA Federpiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,

RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,

RA Gonzalez A., Kremetska I., Kim C., Lenz C., Li J., Liu S.,

RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotska V.S.,

RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN EMBL: U83245; AAC49751.1; -

DR EMBL: AC007258; AAD39318.1; -

DR Mendel, 24148; Arabidopsis 3165; 24148.

DR InterPro: IPR003311; -

DR InterPro: IPR003340; -

DR Pfam: PF02309; AUX_1TA; 1.

DR Pfam: PF02362; B3; 1.

SQ SEQUENCE 665 AA; 73667 MW; 79DD3180C2091401 CRC64;

QY 2 RRRPRPP 10

DB 372 KRPRPPGIP 380

RESULT 29

09VM47 PRELIMINARY; PRT: 2328 AA.

AC 09VM47;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE CG8491 PROTEIN.

GN CG8491.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BRKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Manatlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostel C., Gabrielian A.E., Gary N.S., Galibert W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigwe C.,

RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003515; AAF49103.1; -

DR FLYBase; FBgn0036912; CG8491.

SQ SEQUENCE 2328 AA; 258011 MW; 190B8AD607A45206 CRC64;

QY 1 RRR-----PPPPYLP 11

DB 583 RRRDDYPRPPYEPK 599

Query Match 64.1%; Score 41; DB 5; Length 2328;

Best Local Similarity 52.9%; Pred. No. 3.8e+02;

Matches 9; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

RESULT 30

09GPH4 PRELIMINARY; PRT: 2531 AA.

AC 09GPH4;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE TRANSCRIPTIONAL COACTIVATOR KOHTALO.

GN KTO.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA Treisman J.E.;

RT "Drosophila homologs of the transcriptional coactivation complex in

RT subunits TRAP240 and TRAP230 are required for identical processes in

RT eye-antennal disc development.";

RL Development 0:0-0(2001).

DR EMBL: AF324426; AAG48328.1; -

SQ SEQUENCE 2531 AA; 279500 MW; 68B0CD469EB3A9F5 CRC64;

Query Match 64.1%; Score 41; DB 5; Length 2531;
 Best Local Similarity 52.9%; Pred. No. 4; 1e+02;
 Matches 9; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

OY 1 RRR-----PRPPYLP 11
 DB 586 RRREDQVEPRPPYEPK 602

RESULT 31
 ID C30708 PRELIMINARY; PRT; 38 AA.
 AC C30708;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 GN SPORE COAT PROTEIN (FRAGMENT).
 OS *Bacillus subtilis*.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 FC STRAIN-168;
 RX MEDLINE=98240225; PubMed=9579062;
 RA Rivolta C., Solito B., Lazarevic V., Joris B., Manuel C., Karamata D.;
 RT "A 35.7 kb DNA fragment from the *Bacillus subtilis* chromosome
 containing a putative 12.3 kb operon involved in hexuronate catabolism
 and a perfectly symmetrical hypothetical catabolite-responsive
 element.";
 RL Microbiology 144:877-884 (1998).
 DR EMBL; AF015825; AAC46305.1; -.
 KW Coat protein.
 FT NON_TER 38
 SQ SEQUENCE 38 AA; 4882 MW; 89EF2DF935B66854 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 38;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 RPPPYLP 10
 DB 26 RPPPYLP 33

RESULT 32
 ID 024156 PRELIMINARY; PRT; 129 AA.
 AC 024156;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE GLYCINE-RICH PROTEIN (FRAGMENT).
 GN MRCR15.
 OS *Nicotiana tabacum* (Common tobacco).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 CC Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Fuerstenberg S.I., Buccaglia P.A., Smith A.G.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U74325; AAB18261.1; -.
 DR Mendel; 25642; Nicta; 343; 25642.
 FT NON_TER 1
 SQ SEQUENCE 129 AA; 12251 MW; C6F5A54DE29F7597 CRC64;

Query Match 62.5%; Score 40; DB 10; Length 129;

Best Local Similarity 87.5%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPYLP 9
 DB 28 RRRPPYLP 35

RESULT 33
 ID 006242 PRELIMINARY; PRT; 262 AA.
 AC 006242;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 28.3 KDA PROTEIN.
 GN RV2133C OR MTCY270.35.
 OS *Mycobacterium tuberculosis*.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Horsby T., Jajels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of *Mycobacterium tuberculosis* from the
 complete genome sequence.";
 RL Nature 393:537-544 (1998).
 DR EMBL; 295388; CAB08660.1; -.
 KW Tuberculin; RV2133C; -.
 FT Hypothetical protein
 SQ SEQUENCE 262 AA; 28316 MW; C408B554E3DCDD34 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 262;
 Best Local Similarity 87.5%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RPPPYLP 10
 DB 116 RPPPYLP 123

RESULT 34
 ID 09F345 PRELIMINARY; PRT; 287 AA.
 AC 09F345;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 32.2 KDA PROTEIN.
 GN SC9E12.10C.
 OS *Streptomyces coelicolor*.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kiser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinash H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
KW EMBL: AL391751; CAC05755.1; -.
DR Hypothetical protein
SQ SEQUENCE 287 AA; 32166 MW; 9951067A57D99AE3 CRC64;

Query Match 62.5%; Score 40; DB 2; Length 287;
Best Local Similarity 72.7%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRRPPPLPR 11
DB 67 RRRPPPLPR 77

RESULT 35
ID 041848 PRELIMINARY; PRT; 301 AA.
AC 041848;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PROLIN RICH PROTEIN.
GN zea mays (Maize).
OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-W64A;
RX MEDLINE=92361259; PubMed=1498600;
RA Jose-Estanyol M., Ruiz-Avila L., Puigdomenech P.;
RT "A maize embryo-specific gene encodes a proline-rich and hydrophobic
protein.";
RL Plant Cell 4:413-423(1992).
DR EMBL: X60432; CAA42959.1; -.
DR HSSP: P24337; 1HYP.
DR Mendel: 16222; Zeama;1531;16222.
DR InterPro: IPR000528; -.
DR InterPro: IPR003612; -.
DR Pfam: PF00279; LTP; 1.
DR SMART: SM00499; AAI; 1.
SQ SEQUENCE 301 AA; 31647 MW; 884EB70854D28C2E CRC64;

Query Match 62.5%; Score 40; DB 10; Length 301;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRRPPPLPR 10
DB 83 RRRPPPLPR 90

RESULT 36
ID 081299 PRELIMINARY; PRT; 324 AA.
AC 081299;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE EPOXIDE HYDROLASE (EC 3.3.2.3).

GN T14P8.15 OR AT4G02340.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RT WASHU;
RA "The A. thaliana Genome Sequencing Project.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Katicki J., Elliott G., Cloud J.;
RT "The sequence of A. thaliana T14P8.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Waterston R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BIOTRANSFORMATION ENZYME THAT CATALYZES THE HYDROLYSIS
OF EPOXIDES (ALKENE OXIDES, OXYANES) AND ARENE OXIDES TO LESS
REACTIVE AND MORE WATER SOLUBLE DIHYDRODIOLS BY THE TRANS ADDITION
OF WATER (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: AN EPOXIDE + H(2)O = A GLYCOL (BY SIMILARITY).
CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DR EMBL: AF069298; AAC19281.1; -.
DR EMBL: AL161494; CAB80727.1; -.
DR HSSP: P34914; 1CR6.
DR Mendel: 31939; Arab;1147;31939.
DR InterPro: IPR000073; -.
DR InterPro: IPR000379; -.
DR InterPro: IPR000639; -.
DR InterPro: IPR002106; -.
DR Pfam: PF00561; abhydrolase; 1.
DR PRINTS: PR00412; EPOXHYDRLASE.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
KW Aromatic hydrocarbons catabolism; Detoxification; Hydrolase.
SQ SEQUENCE 324 AA; 36657 MW; F7FDBFC7D6DBA04 CRC64;

Query Match 62.5%; Score 40; DB 10; Length 324;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRRPPPLPR 11
DB 180 RRRPPPLPR 189

RESULT 37
ID 070220 PRELIMINARY; PRT; 400 AA.
AC 070220;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE FORK HEAD TRANSCRIPTION FACTOR.
GN HPHIL OR HPH-1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129; TISSUE-KIDNEY;
RX MEDLINE=98392851; PubMed=9726250;
RA Frank S.; Zoll B.;
RT "Mouse HNF-3/fork head homolog-1-like gene: structure, chromosomal
   location, and expression in adult and embryonic kidney.";
RL DNA Cell Biol. 17:679-688(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129; TISSUE-KIDNEY;
RA Frank S.; Zoll B.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-129; TISSUE-KIDNEY;
RA Pasche B.; Bieller A.; Zoll B.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF010405; AAC12973.2; -.
DR HSSP; Q63245; 2HFH.
DR MGD; MGI:1298228; Hfhl1.
DR InterPro: IPR001766; -.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
DR SMART; SM00339; FH; 1.
SQ SEQUENCE 400 AA; 41342 MW; E6C3B5F53BC42B CRC64;

OY 2 RRRPPY 8
Db 112 RRRPPY 118

Query Match 62.5%; Score 40; DB 11; Length 400;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 38
O9JUN7 PRELIMINARY; PRT; 400 AA.
AC O9JUN7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE WINGED HELIX/FORKHEAD TRANSCRIPTION FACTOR HFH1.
GN HFH1.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong H.-K.; Noveroske J.K.; Justice M.J.; Chakravarti A.;
RT "The winged helix/forkhead transcription factor Hfh1 gene regulates
   hair keratinization in satin mice.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF153193; AAF73234.1; -.
DR InterPro: IPR001766; -.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
DR SMART; SM00339; FH; 1.
SQ SEQUENCE 400 AA; 41444 MW; 99318026C8E2AD19 CRC64;

Query Match 62.5%; Score 40; DB 11; Length 400;

```

```

Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPPY 8
Db 112 RRRPPY 118

RESULT 39
O9JUN7 PRELIMINARY; PRT; 400 AA.
AC O9JUN7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HFH1.
GN HFH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RA Hong H.-K.; Noveroske J.K.; Justice M.J.; Chakravarti A.;
RT "The winged helix/forkhead transcription factor Hfh1 gene regulates
   hair keratinization in satin mice.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154426; AAF74524.1; -.
DR InterPro: IPR001766; -.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
DR SMART; SM00339; FH; 1.
SQ SEQUENCE 400 AA; 41368 MW; F286AEFE81D4759B CRC64;

OY 2 RRRPPY 8
Db 112 RRRPPY 118

RESULT 40
O9NS06 PRELIMINARY; PRT; 402 AA.
AC O9NS06;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE WINGED HELIX/FORKHEAD TRANSCRIPTION FACTOR.
GN HFH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong H.-K.; Noveroske J.K.; Justice M.J.; Chakravarti A.;
RT "The winged helix/forkhead transcription factor Hfh1 gene regulates
   hair keratinization in satin mice.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF153341; AAF75586.1; -.
DR InterPro: IPR001766; -.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.

```

DR PROSITE; PS50039; FORK_HEAD_3; 1.
 DR SMART; SM00339; FH; 1.
 SQ SEQUENCE 402 AA; 41485 MW; FD2EEF9D848E77 CRC64;

Query Match 62.5%; Score 40; DB 4; Length 402;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 RRPPEPY 8
 |||:|
 Db 115 RRPPEPY 121

RESULT 41
 09VNA6 PRELIMINARY; PRT; 596 AA.
 AC 09VNA6; PRELIMINARY; PRT; 596 AA.
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CG16708 PROTEIN.
 GN CG16708.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_Taxid=7227;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amandlis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt A.D., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Abghyani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Berens P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 Burks K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacibel J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster.";
 RT Science 287:2185-2195(2000).
 RL EMBL; AE003603; AAF52040.1;
 DR Flybase; FBgn0037315; CG16708.
 DR InterPro; IPR001206; -.

DR Pfam; PF00781; DAGKC; 1.
 DR SMART; SM00046; DAGKC; 1.
 SQ SEQUENCE 596 AA; 65622 MW; 91410AC18AAC4660 CRC64;

Query Match 62.5%; Score 40; DB 5; Length 596;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 5 RPYLPR 11
 |||:|
 Db 294 RPYLPR 300

RESULT 42
 09HV05 PRELIMINARY; PRT; 600 AA.
 AC 09HV05; PRELIMINARY; PRT; 600 AA.
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL PROTEIN PA4517.
 GN PA4517.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 CX NCBI_Taxid=287;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Garber R.J., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL EMBL; AE004865; AAC07905.1;
 DR InterPro; IPR003371;
 DR Pfam; PF02418; DUF146; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 600 AA; 67957 MW; DB8E1A3ED793F7EE CRC64;

Query Match 62.5%; Score 40; DB 2; Length 600;
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 RRPPEPYLPR 11
 |||:|
 Db 141 RRPPEPYLPR 150

RESULT 43
 091889 PRELIMINARY; PRT; 693 AA.
 AC 091889; PRELIMINARY; PRT; 693 AA.
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE UBQUITIN-LIKE FUSION PROTEIN.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 CX NCBI_Taxid=8355;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=93292985; PubMed=8390387;
 RA Linmen J.M., Bailey C.P., Weeks D.L.;

```

RT      "Two related localized mRNAs from Xenopus laevis encode ubiquitin-like
RT      fusion proteins.";
FL      Gene 128:181-188(1993).
DR      EMBL; L08474; AAB04151.1; -.
DR      HSSP; P02248; I0B1.
DR      InterPro; IPR000058; -.
DR      InterPro; IPR000626; -.
DR      Pfam; PF00240; ubiqutin_1.
DR      Pfam; PF01428; zf-ANI; 1.
DR      PRINTS; PR00348; UBIQUITIN.
DR      PROSITE; PS50053; UBIQUITIN_2; 1.
DR      SMART; SM00154; Znf_AN1; 1.
SQ      SEQUENCE 693 AA; 76862 MW; 42ECFAD622697BA0 CRC64;

Query Match
Best Local Similarity 62.5%; Score 40; DB 13; Length 693;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy      3 RRRPPYLP 11
Db      229 KRRPVVPR 237

RESULT  44
Oy      091890
AC      PRELIMINARY; PRT; 701 AA.
DR      01-NOV-1996 (TREMBlrel. 01, Created)
DR      01-NOV-1996 (TREMBlrel. 01, last sequence update)
DR      01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE      UBIQUITIN-LIKE FUSION PROTEIN.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=83355;
RN      [1]
RR      SEQUENCE FROM N.A.
RX      MEDLINE-93292985; PubMed-8390387;
RA      Linmen J.M., Bailey C.P., Weeks D.L.;
RT      "Two related localized mRNAs from Xenopus laevis encode ubiquitin-like
RT      fusion proteins.";
RL      Gene 128:181-188(1993).
DR      EMBL; L08475; AAB49979.1; -.
DR      HSSP; P02248; I1B8.
DR      InterPro; IPR000058; -.
DR      InterPro; IPR000626; -.
DR      Pfam; PF00240; ubiqutin_1.
DR      Pfam; PF01428; zf-ANI; 1.
DR      PRINTS; PR00348; UBIQUITIN.
DR      PROSITE; PS50053; UBIQUITIN_2; 1.
DR      SMART; SM00154; Znf_AN1; 1.
SQ      SEQUENCE 701 AA; 78581 MW; E5D616C032CE4139 CRC64;

Query Match
Best Local Similarity 62.5%; Score 40; DB 13; Length 701;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy      3 RRRPPYLP 11
Db      228 KRRPVVPR 236

RESULT  45
Oy      091905
AC      PRELIMINARY; PRT; 134 AA.
DR      01-NOV-1999 (TREMBlrel. 12, Created)
DR      01-NOV-1999 (TREMBlrel. 12, last sequence update)
DR      01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE      HYPOTHETICAL 14.7 KDA PROTEIN APE2233.

```

```

GN      APE2233.
OS      Aeropyrum pernix.
OC      Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC      Aeropyrum.
OX      NCBI_TaxID=56636;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-K1.
RX      MEDLINE-99310339; PubMed-10382966;
RA      Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA      Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA      Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA      Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA      Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA      Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT      "Complete genome sequence of an aerobic hyper-thermophilic
RT      crenarchaeon, Aeropyrum pernix K1.";
RL      DNA Res. 6:83-101(1999).
DR      EMBL; AP000063; BAA81245.1; -.
DR      Hypothetical protein.
SQ      SEQUENCE 134 AA; 14662 MW; D02474F86FB71AB CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 1; Length 134;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      1 RRRPPYLP 10
Db      3 RRRPAPYSP 12

```

Search completed: September 24, 2001, 10:07:42
 Job time: 142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 24, 2001, 10:08:36 ; Search time 44.37 Seconds
(without alignments)
10.931 Million cell updates/sec

Title: US-09-276-868-5
Perfect score: 48
Sequence: 1 RRRPRPPY 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSB8/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDSB8/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDSB8/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDSB8/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDSB8/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDSB8/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDSB8/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDSB8/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDSB8/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDSB8/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDSB8/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDSB8/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDSB8/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDSB8/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDSB8/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDSB8/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDSB8/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDSB8/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDSB8/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDSB8/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDSB8/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDSB8/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	8	21	AAB26887
2	48	100.0	11	21	AAB26886
3	48	100.0	14	17	AAW01450
4	48	100.0	14	19	AAW5725
5	48	100.0	15	21	AAB26885
6	48	100.0	19	17	AAW01452
7	48	100.0	26	17	AAW01447
8	48	100.0	26	19	AAW5723
9	48	100.0	39	14	AAW30491
10	48	100.0	39	17	AAW01446
11	48	100.0	39	17	AAW94446

12	48	100.0	39	17	AAW99121
13	48	100.0	39	19	AAW57522
14	48	100.0	39	21	AAB26888
15	48	100.0	44	22	AAW51194
16	41	85.4	106	20	AAW29263
17	41	85.4	272	21	AAW57119
18	41	85.4	548	18	AAW40306
19	37	77.1	288	21	AAW13899
20	37	77.1	322	21	AAW13898
21	37	77.1	340	21	AAW06371
22	37	77.1	398	21	AAW06370
23	37	77.1	412	21	AAW06369
24	36	75.0	21	19	AAW21446
25	36	75.0	30	14	AAW39869
26	36	75.0	39	20	AAW74025
27	36	75.0	41	16	AAW79134
28	36	75.0	46	14	AAW50053
29	36	75.0	70	21	AAW10803
30	36	75.0	73	21	AAW07230
31	36	75.0	76	21	AAW10802
32	36	75.0	155	21	AAW25254
33	36	75.0	166	21	AAW26907
34	36	75.0	166	21	AAW27133
35	36	75.0	169	14	AAW50051
36	36	75.0	179	14	AAW50056
37	36	75.0	235	20	AAW73974
38	36	75.0	252	13	AAW21708
39	36	75.0	258	14	AAW50049
40	36	75.0	263	13	AAW21706
41	36	75.0	294	16	AAW78729
42	36	75.0	294	18	AAW26589
43	36	75.0	338	13	AAW21707
44	36	75.0	388	16	AAW78734
45	36	75.0	388	18	AAW26592

ALIGNMENTS

RESULT 1
ID AAB26887 standard; peptide; 8 AA.
XX AAB26887;
AC
DT 01-FEB-2001 (first entry)
XX
DE PR-39 derived angiogenesis regulatory peptide 3.
XX
KW Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
KW myocardial ischaemia; proteasome.
XX
OS Synthetic.
XX
PN WO200057895-A1.
XX
PD 05-OCT-2000.
XX
PF 16-MAR-2000; 2000WO-US07050.
XX
PR 26-MAR-1999; 99US-0276868.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Simons M, Gao Y;
XX
DR WPI; 2000-628319/60.
XX
PT Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
PT infarction, by administering a PR-39 oligopeptide that regulates
XX enzymatic activity of proteasomes
XX
PS Claim 14; Page 41; Sipp; English.

XX This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
 CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents a PR-39 derived peptide which interacts with the proteasome
 CC and can be used in the method of the invention.

XX
 SQ Sequence 8 AA;

Query Match 100.0%; Score 48; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
 |||||
 DB 1 rrrprpy 8

RESULT 2

AAB26886
 ID AAB26886 standard; peptide; 11 AA.

AC AAB26886;

DT 01-FEB-2001 (first entry)

PR-39 derived angiogenesis regulatory peptide 2.

Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
 KM myocardial ischemia; proteasome.

OS Synthetic.

PN WO200057895-A1.

XX 05-OCT-2000.

PF 16-MAR-2000; 2000WO-US07050.

PR 26-MAR-1999; 99US-0276868.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Simons M, Gao Y;

DR WPI: 2000-628319/60.

XX Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
 PT infarction, by administering a PR-39 oligopeptide that regulates
 PT enzymatic activity of proteasomes

PS Claim 13; Page 41; 51pp; English.

XX This invention relates to a method for the stimulation of angiogenesis in
 CC situ within a targeted collection of viable cells. The method comprises
 CC introducing, into the cytoplasm, at least 1 member of the PR-39
 CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
 CC Part of the proteolytic activity of the proteasomes is selectively
 CC altered so as to stimulate angiogenesis. The method is used to induce
 CC angiogenesis in tissue that has suffered anoxia or infarction,
 CC e.g. myocardial infarction or chronic myocardial ischemia, and also to
 CC study the mechanisms that control angiogenesis. The present sequence
 CC represents a PR-39 derived peptide which interacts with the proteasome
 CC and can be used in the method of the invention.

XX
 SQ Sequence 11 AA;

Query Match 100.0%; Score 48; DB 21; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
 |||||
 DB 1 rrrprpy 8

RESULT 3

AAW01450
 ID AAW01450 standard; peptide; 14 AA.

AC AAW01450;

DT 18-JUN-1997 (first entry)

DE Leukocyte O2- production inhibitor peptide PR14.

XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
 KM antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KM DNA synthesis; protein synthesis; inhibitor; syndecan expression;
 KM mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KM tissue damage; oxygen radical; inflammatory disease; therapy.

OS Synthetic.

PN WO9632129-A1.

XX 17-OCT-1996.

PF 10-APR-1996; 96WO-US04674.

PR 10-APR-1995; 95US-0419066.

XX (UNITV) UNIV KANSAS STATE RES FOUND.

PA Blecha F, Shi J;

DR WPI: 1996-476842/47.

XX Inhibition of leukocyte superoxide anion prodn. and attraction of
 PT leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 XX
 PS Claim 12; Page 28; 45pp; English.

XX AAW01447-W01454 represent fragments of the proline-arginine rich
 CC antimicrobial peptide PR39 (see AAW01446). The PR39 sequence was first
 CC isolated from porcine small intestine, and has also been identified in
 CC human and porcine neutrophils. PR39 kills bacteria by interfering with
 CC DNA and/or protein synthesis. PR39 also induces syndecan expression on
 CC mesenchymal cells. Syndecans are important in wound repair, showing that
 CC PR39 can be used in wound repair, as well as in antibacterial agents.
 CC These sequences, and PR39, can be used in the method of the invention.
 CC The method of the invention is for inhibiting leukocyte superoxide anion
 CC (O2-) production. The method comprises administering to a leukocyte a
 CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
 CC production. The peptides can be used as medicaments for fighting
 CC infection by attracting leukocytes to a wound site and restricting
 CC tissue damage at the wound site caused by excessive oxygen radicals
 CC produced by these leukocytes. They can also be used to develop products
 CC for treating inflammatory disease states.

XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 48; DB 17; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.33; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

OY 1 RRRPRPY 8

xx Blecha F, Ross CR, Shi J;
 xx WPI: 1998-495359/42.
 xx Reduction of reperfusion injury in temporarily occluded blood
 xx vessels - by administration of a peptide which is rich in proline
 xx or arginine residues
 xx
 xx Claim 3: Page 14-15; 35pp; English.
 xx
 xx Sequences AAW5722-W75732 are proline/arginine rich peptides that upon
 xx administration into a mammal's bloodstream reduce reperfusion injury
 xx (production of reactive oxygen species, neutrophil adherence to
 xx endothelium, and extravasation of neutrophils). These peptides have two
 xx requirements: they contain the consensus sequence PXXP, where P is a
 xx proline residue and X is any amino acid residue, which has been found to
 xx inhibit superoxide production, and secondly they have arginine residues
 xx adjacent to these motifs, required for effective inhibition. It was
 xx established by structural and function analysis that a peptide should
 xx ideally contain 4 or 6 of these motifs, and that inhibitory activity is
 xx correlated with the increase of length of peptides. The effectiveness
 xx of these peptides was determined by investigating the production of the
 xx neutrophil superoxide anion, and also the inhibition of neutrophil
 xx chemotaxis. From this, it was found that all of the peptides inhibited
 xx NADPH oxidase to some extent. All of the peptides also inhibit
 xx neutrophil oxidase activity. PR-39 is believed, to be the most potent
 xx endogenous down regulator of NADPH oxidase yet discovered, and from the
 xx data produced, it can be suggested to be involved in eliminating or
 xx reducing the reperfusion injury induced adhesion and extraction of
 xx neutrophils. The peptides are also useful in connection with surgical
 xx procedures such as coronary bypass and organ transplantation surgery.
 xx
 xx Sequence 26 AA;

Query Match 100.0%; Score 48; DB 19; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
 |||||
 Db 1 rrrprppy 8
 RESULT 9
 AAR30491
 ID AAR30491 standard; peptide; 39 AA.
 AC AAR30491;
 DT 12-MAY-1993 (first entry)
 XX
 XX Antibacterial peptide.
 DE
 XX Pig; small intestine; endocrine; gram negative; bacteria; therapeutic;
 KW veterinary medicine; prophylactic.
 XX
 XX Sus scrofa domestica.
 OS
 XX WO9222578-A.
 XX
 XX 23-DEC-1992.
 PD
 XX
 XX 10-JUN-1992; 92WO-SE00394.
 PF
 XX
 XX 14-JUN-1991; 91SE-0001838.
 PR
 XX (BOMA/) BOMAN H G.
 PA (JOER/) JOERNVALL H.
 RA (LEJ/) LEE J.
 PA (MUTT/) MUTT V.
 XX

PI Roman HG, Joernvall H, Lee J, Mutt V;
 XX WPI: 1993-018080/02.
 DR
 XX New anti-bacterial polypeptide - active against Gram negative
 PT bacteria
 PS
 XX Claim 1: Page 10; 15pp; English.
 XX
 XX This peptide was isolated from the small intestine of a pig. The
 XX small intestine is an important endocrine organ and many
 XX physiologically active peptides have been isolated from it. This
 XX peptide inhibits the growth of, and may kill, bacteria, pref. gram
 XX negative bacteria. This peptide or its functional derivatives may be
 XX used in human or veterinary medicine for therapeutic or prophylactic
 CC use.
 CC
 CC Sequence 39 AA;

Query Match 100.0%; Score 48; DB 14; Length 39;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
 |||||
 Db 1 rrrprppy 8

RESULT 10
 AAW01446
 ID AAW01446 standard; peptide; 39 AA.
 AC AAW01446;
 DT
 XX
 XX 18-JUN-1997 (first entry)
 DT
 XX
 XX Leukocyte O2- production inhibitor peptide PR39.
 DE
 XX
 XX Inhibitor; leukocyte O2- production; proline-arginine rich peptide; pig;
 KW antimicrobial peptide; small intestine; human; neutrophil; bacteria;
 KW DNA synthesis; protein synthesis; inhibitor; syndecin expression;
 KW mesenchymal cell; wound repair; superoxide anion; infection; leukocyte;
 KW tissue damage; oxygen radical; inflammatory disease; therapy.
 XX
 XX Synthetic.
 OS
 XX WO9632129-A1.
 PN
 XX 17-OCT-1996.
 PD
 XX
 XX 10-APR-1996; 96WO-US04674.
 PF
 XX
 XX 10-APR-1995; 95US-0419066.
 PR
 XX (UNIV) UNIV KANSAS STATE RES FOUND.
 PA
 XX Blecha F, Shi J;
 PI
 XX WPI: 1996-476842/47.
 DR
 XX
 XX Inhibition of leukocyte superoxide anion prodn. and attraction of
 PT leukocytes - using peptide(s) partic. based on antimicrobial PR-39
 XX
 XX Claim 2; Page 26; 45pp; English.
 PS

This sequence represents the proline-arginine rich antimicrobial peptide PR39. The PR39 sequence was first isolated from porcine small intestine, and has also been identified in human and porcine neutrophils. PR39 kills bacteria by interfering with DNA and/or protein synthesis. PR39 also induces syndecin expression on mesenchymal cells. Syndecins are important in wound repair, showing that PR39 can be used in wound repair, as well as in antibacterial agents. This sequence, and the fragments of

CC It shown in AAW01447-W01454, can be used in the method of the invention.
CC The method of the invention is for inhibiting leukocyte superoxide anion
CC (O2-) production. The method comprises administering to a leukocyte a
CC peptide (such as this sequence) capable of inhibiting leukocyte O2-
CC production. The peptides can be used as medicaments for fighting
CC infection by attracting leukocytes to a wound site and restricting tissue
CC damage at the wound site caused by excessive oxygen radicals produced by
CC these leukocytes. They can also be used to develop products for treating
CC inflammatory disease states.

XX Sequence 39 AA:

Query Match 100.0%; Score 48; DB 17; Length 39;

Best Local Similarity 100.0%; Pred. No. 0.8; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
|||||

DB 1 rrrprppy 8

RESULT 11

AAR94446 AAR94446 standard; peptide; 39 AA.

AC AAR94446;

DT 05-NOV-1996 (first entry)

XX Synducin peptide (PR-39) induces syndecan expression.

XX Synducin: induction; expression; syndecan-1; syndecan-4; surface;
XX mesenchymal cell; fibroblast; epithelial; PR-39; treatment; stasis;
XX decubitus; ulcers; keloids; skin burns; ischemic tissues;
XX hypercoagulation states; prevention; tumour metastasis; restenosis;
XX inhibition; angiogenesis; proliferation; endothelial.

XX Synthetic.

XX WO9609322-A2.

XX 28-MAR-1996.

XX 22-SEP-1995; 95WO-US12080.

XX 22-SEP-1994; 94US-0310722.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Bernfield M, Gallo RL;

XX WPI; 1996-188401/19.

XX Modulating mesenchymal interaction by administration of synducin
XX used in the treatment of wounds, tumours, restenosis, etc

XX Claim 4; Page 26; 34pp; English.

XX The present peptide is a synducin, which induces the expression of
XX syndecan-1 and syndecan-4 on the surface of mesenchymal cells, esp.
XX fibroblasts and epithelial cells. The 36 N-terminal amino acids of
XX the peptide were found to be identical to the 36 N-terminal amino
XX acids of PR-39, a Pro and Arg rich antibacterial peptide previously
XX found in porcine intestine (WO9222578). Synducins may be used in
XX the treatment of stasis and decubitus ulcers, keloids, skin burns,
XX ischemic tissues and hypercoagulation states, prevention of tumour
XX metastasis, restenosis inhibition and endothelial cell angiogenesis
XX and proliferation induction.

XX Human microvascular endothelial cells were assayed for syndecan-4
XX expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
XX the present peptide (10 microm) or a blank, to give respective
XX cell surface syndecan-4 values (mD/m in) of approx. 1.75, 1.70,

CC 1.80 and 0.95.

XX Sequence 39 AA:

Query Match 100.0%; Score 48; DB 17; Length 39;

Best Local Similarity 100.0%; Pred. No. 0.8; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
|||||

DB 1 rrrprppy 8

RESULT 12

AAR9121 AAR9121 standard; peptide; 39 AA.

AC AAR9121;

DT 28-OCT-1996 (first entry)

XX Magainin-derived antimicrobial STD-inhibiting peptide, MST-1312.

XX STD; sexually transmitted disease; HIV; human immunodeficiency virus;
XX herpes simplex virus; HSV; Neisseria gonorrhoeae; Candida; Chlamydia;
XX magainin; antimicrobial; squalamine.

XX Synthetic.

XX Key Location/Qualifiers

XX Modified-site 39

XX /note= "amidated"

XX WO9608270-A2.

XX 21-MAR-1996.

XX 13-SEP-1995; 95WO-US11675.

XX 13-SEP-1994; 94US-0305475.

XX (MAGA-) MAGAININ PHARM INC.

XX Bedi G, Jacob L, Williams T, Zasloff M;

XX WPI; 1996-179725/18.

XX Inhibiting sexually transmitted disease e.g. HIV or herpes simplex -
XX by administering magainin antimicrobial or squalamine cpd. to
XX inhibit transmission

XX Example 1; Page 32; 60pp; English.

XX AAR9116-R9123 are antimicrobial, magainin-analogue peptides that may
XX be used to treat sexually transmitted diseases (STDs) caused by
XX Chlamydia, HIV, herpes simplex virus, Neisseria gonorrhoeae or
XX Candida infection. The peptides inhibit STDs by either killing the
XX infectious organism, impeding the infection mechanism or
XX interrupting the replication cycle of the organism. Squalamine (an
XX amino-steroid host defence molecule of the dog fish shark Squalus
XX acanthias) and Pgla (a frog antimicrobial peptide) analogues may
XX also be useful in inhibiting STD infection and transmission.

XX Sequence 39 AA:

Query Match 100.0%; Score 48; DB 17; Length 39;

Best Local Similarity 100.0%; Pred. No. 0.8; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
|||||

Db 1 rrrrrpppy 8

RESULT 13

ID AAW75722

AAW75722 standard; peptide: 39 AA.

AC AAW75722;

ET 19-NOV-1998 (first entry)

CE Proline/Arginine rich peptide PR-39.

KW Proline: arginine; peptide: reperfusion injury; neutrophil; endothelium;
KW superoxide; neutrophil superoxide anion; chemotaxis; NADPH oxidase;
KW coronary bypass; organ transplantation surgery.

CS Synthetic.

FN WO9835690-A1.

PD 20-AUG-1998.

PF 17-FEB-1998; 98WO-US03207.

PR 16-FEB-1998; 98US-0024975.

PR 18-FEB-1997; 97US-0802306.

PA (UNITV) UNIV KANSAS STATE RES FOUND.

PI Blecha F, Ross CR, Shi J;

WPI: 1998-495359/42.

PT Reduction of reperfusion injury in temporarily occluded blood
PT vessels - by administration of a peptide which is rich in proline
PT or arginine residues

PS Claim 3: Page 14; 35pp; English.

CC Sequences AAW75722-W75722 are proline/arginine rich peptides that upon
CC administration into a mammal's bloodstream reduce reperfusion injury
CC (production of reactive oxygen species, neutrophil adherence to
CC endothelium, and extravasation of neutrophils). These peptides have two
CC requirements: they contain the consensus sequence PXXP, where P is a
CC proline residue and X is any amino acid residue, which has been found to
CC inhibit superoxide production, and secondly they have arginine residues
CC adjacent to these motifs, required for effective inhibition. It was
CC established by structural and function analysis that a peptide should
CC ideally contain 4 or 6 of these motifs, and that inhibitory activity is
CC correlated with the increase of length of peptides. The effectiveness
CC of these peptides was determined by investigating the production of the
CC neutrophil superoxide anion, and also the inhibition of neutrophil
CC chemotaxis. From this, it was found that all of the peptides inhibited
CC NADPH oxidase to some extent. All of the peptides also inhibit
CC neutrophil oxidase activity. PR-39 is believed, to be the most potent
CC endogenous down regulator of NADPH oxidase yet discovered, and from the
CC data produced, it can be suggested to be involved in eliminating or
CC reducing the reperfusion injury induced adhesion and extraction of
CC neutrophils. The peptides are also useful in connection with surgical
CC procedures such as coronary bypass and organ transplantation surgery.

XX Sequence 39 AA;

Query Match 100.0%; Score 48; DB 19; Length 39;

Best Local Similarity 100.0%; Pred No. 0.8; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRRPPPY 8

DB 1 rrrrrpppy 8

RESULT 14

ID AAB26888

AAB26888 standard; peptide: 39 AA.

AC AAB26888;

DT 01-FEB-2001 (first entry)

DE PR-39 peptide used in angiogenesis control.

KW Angiogenesis; stimulation; PR-39; anoxia; myocardial infarction;
KW myocardial ischaemia; proteasome.

OS Synthetic.

PN WO200057895-A1.

PD 05-OCT-2000.

PF 16-MAR-2000; 2000MO-US07050.

PR 26-MAR-1999; 99US-0276868.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Simons M, Gao Y;

WPI: 2000-628319/60.

PT Stimulating angiogenesis in situ, useful e.g. for treating anoxia and
PT infarction, by administering a PR-39 oligopeptide that regulates
PT enzymatic activity of proteasomes

PS Disclosure; Page 21; 51pp; English.

CC This invention relates to a method for the stimulation of angiogenesis in
CC situ within a targeted collection of viable cells. The method comprises
CC introducing, into the cytoplasm, at least 1 member of the PR-39
CC oligopeptide collective, which interacts with cytoplasmic proteasomes.
CC Part of the proteolytic activity of the proteasomes is selectively
CC altered so as to stimulate angiogenesis. The method is used to induce
CC angiogenesis in tissue that has suffered anoxia or infarction,
CC e.g. myocardial infarction or chronic myocardial ischaemia, and also to
CC study the mechanisms that control angiogenesis. The present sequence
CC represents the PR-39 peptide from which peptide used in the method of
CC the invention are derived.

SQ Sequence 39 AA;

Query Match 100.0%; Score 48; DB 21; Length 39;

Best Local Similarity 100.0%; Pred. No. 0.8; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0;

OY 1 RRRRRPPPY 8

DB 1 rrrrrpppy 8

RESULT 15

ID AAB51194

AAB51194 standard; Peptide: 44 AA.

AC AAB51194;

DT 22-MAR-2001 (first entry)

DE E. coli AMP gene PR39 amino acid sequence.
DE Escherichia coli; E. coli; AMP gene; anti-microbial peptide;
KW screening; preservation; food; feed; paint formulation; detergent;
KW cosmetic; medical device; prosthetic implant; disinfectant;

```
KW microbial infection; tumour.
XX
XX Escherichia coli.
XX
XX PN W0200073433-A1.
XX
XX PD 07-DEC-2000.
XX
XX PF 29-MAY-2000; 2000WO-DK00287.
XX
XX PR 31-MAY-1999; 99DK-0000766.
XX
XX PA (NOVO ) NOVO NORDISK AS.
XX
XX PI Kristensen H;
XX
XX DR WPI; 2001-070965/08.
XX
XX PT Screening for nucleotide sequences encoding antimicrobial peptides by
XX generating peptide libraries in microorganisms, inducing expression of
XX peptides, selecting cells and recovering nucleotide sequences from
XX cells
XX
XX PS Example 1; Page 34; 59pp; English.
XX
XX CC The present invention describes a plasmid which is ligated with a pool
XX of nucleotide sequences (NT) linked to an inducible promoter, to express
XX a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
XX optionally linked to a signal (P)), transformed into host cells and
XX cultured in presence of an inducer to induce expression of the NT. A
XX method of screening (1) a pool of nucleotide sequences to select a
XX nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
XX with the pool of NT; (b) transforming host cells which are sensitive to
XX the peptide with the ligated plasmids; (c) screening the transformed
XX cells to select viable cells; (d) cultivating the viable cells in the
XX presence of an inducer to induce expression of NT; (e) selecting cells
XX according to the effect of the inducer on cell growth; and (f) recovering
XX NT encoding the peptide from the selected cells. (1) is useful for
XX screening a pool of nucleotide sequences to select a nucleotide sequence
XX encoding a peptide which is an antimicrobial peptide or an antimicrobial
XX enzyme active on bacteria and for finding and preparing a composition for
XX treatment of human or animal. The antimicrobial peptide obtained using
XX (1) may be employed in preservation of e.g. food/feed, paint
XX formulations, detergents, cosmetics, medical devices such as prosthetic
XX implants and also to disinfect and/or kill microbial cells on an object
XX e.g. as an disinfectant for the treatment of biofilm. The peptides are
XX useful for treating microbial infections and/or tumours. Peptides with
XX improved bio-activity can be developed using (1). The peptides have no
XX negative effect on normal mammalian and/or eukaryotic cells. The present
XX sequence represents an AMP (antimicrobial peptide) gene amino acid
XX sequence, which is used in an example from the present invention.
XX
XX SO Sequence 44 AA.

Query Match 100.0%; Score 48; DB 22; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
   |||||||
DB 3 rrrrpppy 10

RESULT 16
ID AAY29263 standard; Protein; 106 AA.
XX
XX AC AAY29263;
XX
XX DT 25-OCT-1999 (first entry)
XX
XX DE Amino acid sequence of a virulence factor encoded by ORF41598.
```

```
XX
XX KW Human pathogen; virulence polypeptide; virulence factor;
XX pathogenic infection; Pseudomonas aeruginosa infection.
XX
XX OS Pseudomonas aeruginosa.
XX
XX PN W0927129-A1.
XX
XX PD 03-JUN-1999.
XX
XX PF 25-NOV-1998; 98WO-US25247.
XX
XX PR 25-NOV-1997; 97US-0066517.
XX
XX PA (GENO ) GEN HOSPITAL CORP.
XX
XX PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;
XX PI Rahme LG, Tan M, Tsongalis J;
XX
XX DR WPI; 1999-357851/30.
XX
XX PT Virulence factors useful in developing disease treatments
XX
XX PS Disclosure: Fig 4; 228pp; English.
XX
XX CC The present sequence represents a Pseudomonas aeruginosa polypeptide
XX sequence. P. aeruginosa is an opportunistic human pathogen present in
XX soil water and plants. The specification describes virulence polypeptides
XX and nucleic acid sequence encoding such polypeptides. These sequences
XX can be used to identify a compound which is capable of decreasing the
XX expression of a pathogenic virulence factor. Compounds that inhibit
XX the expression or activity of virulence factor polypeptides can be
XX used to treat pathogenic infections, especially where the infection
XX is a P. aeruginosa infection.
XX note: the sequences given in the specification were poorly legible, and
XX in some instances assumptions were made as to the identity of the
XX residue; it is therefore possible that the sequence given below is
XX not entirely correct.
XX
XX SO Sequence 106 AA;

Query Match 85.4%; Score 41; DB 20; Length 106;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPP 7
   |||||||
DB 27 rrrrpp 33

RESULT 17
ID AAB57119 standard; Protein; 272 AA.
XX
XX AC AAB57119;
XX
XX DT 13-MAR-2001 (first entry)
XX
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1697.
XX
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
XX vulnerrary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.
XX
XX OS Homo sapiens.
XX
XX PN W0200055174-A1.
XX
XX PD 21-SEP-2000.
```

```
XX 08-MAR-2000; 2000MO-US05988.
FF
XX
PR 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
ER WPI: 2000-587513/55.
XX N-PSDB; AAF16322.
XX
PR Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX Claim 11; Page 2169-2170; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioprotective, immunomodulatory, muscular, vulnerrary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 272 AA:

Query Match 85.4%; Score 41; DB 21; Length 272;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
   |||||
DB 222 rrrprpp 228

RESULT 18
AAM40306
ID AAM40306 standard; Protein; 548 AA.
XX
XX AAM40306;
XX
XX 07-JUL-1998 (first entry)
XX
DE Parapox virus protein kinase.
XX
XX Recombinant parapox virus; PPV; vaccine; diagnostic; protection;
KM non-pathogen-specific immunostimulation; response; protein kinase;
KM vascular endothelial growth factor; VEGF.
XX
XX Parapox virus.
OS
XX DE19639601-A1.
XX
XX 04-SEP-1997.
XX
XX 26-SEP-1996; 96DE-1039601.
XX
XX 28-FEB-1996; 96DE-1007458.
XX
XX (FARB ) BAYER AG.
XX
XX Bueltner M, Rziha H, Schmeer N, Strube W;
```

```
XX WPI: 1997-436729/41.
DR
DR N-PSDB; AAT76874.
XX
XX Parapox virus strain D1701 HindIII fragment I, related recombinant
PT viruses and expression plasmids - useful in vaccines, for expression
PT of foreign DNA and for diagnosis
XX
PS Claim 28; Page -: 48pp; German.
XX
XX This sequence is protein kinase derived from a HindIII fragment from
CC Parapox virus (PPV) strain D1701. Recombinant PPV, which may contain
CC foreign DNA, are used in vaccines, to protect against the source of the
CC foreign DNA, to produce a non-pathogen-specific immunostimulation or for
CC expression of foreign DNA. The plasmids can be used to express
CC PPV-specific gene sequences and in diagnostic reagents. DNA sequences
CC that can be inserted (preferably 0.1-15 kb) are typically from herpes
CC viruses, foot and mouth disease virus, influenza virus, Pasteurella,
CC Salmonella, Chlamydia, Toxoplasma, Dirofilaria etc. The VEGF promoter can
CC be used to express DNA in PPV. PPV are naturally immunostimulatory so
CC vaccines based on them provide both a quickly developed non-specific
CC protection and a long-lasting pathogen specific response.
CC This sequence is not represented in the specification but is claimed.
XX
SQ Sequence 548 AA:

Query Match 85.4%; Score 41; DB 18; Length 548;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
   |||||
DB 523 rrrprpp 529

RESULT 19
AAG13899
ID AAG13899 standard; Protein; 288 AA.
XX
XX AAG13899;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13560.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
```

PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134253.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
<hr/>		
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148688.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0156559.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159325.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 77.1%; Score 37; DB 21; Length 288;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RRRPP 7
Db 51 rrrppp 57

RESULT 20
AAG13898
ID AAG13898 standard; Protein; 322 AA.
XX AAG13898;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13559.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hydridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 77.1%; Score 37; DB 21; Length 322;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPP 7
Db 85 rrrppp 91

RESULT 21
ID AAG06371 standard; Protein; 340 AA.
AC AAG06371;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3120.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160761.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 77.1%; Score 37; DB 21; Length 340;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 RRRPDP 7
Db 103 rrrpdp 109
RESULT - 22
AAG06370
ID AAG06370 standard; Protein; 398 AA.
AC AAG06370;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3119.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

05-AUG-1999;	9905-0147260.
PR 06-AUG-1999;	9905-0147303.
PR 06-AUG-1999;	9905-0147416.
PR 09-AUG-1999;	9905-0147493.
PR 09-AUG-1999;	9905-0147935.
PR 10-AUG-1999;	9905-0148171.
PR 11-AUG-1999;	9905-0148319.
PR 12-AUG-1999;	9905-0148341.
PR 13-AUG-1999;	9905-0148565.
PR 13-AUG-1999;	9905-0148684.
PR 16-AUG-1999;	9905-0149368.
PR 17-AUG-1999;	9905-0149375.
PR 18-AUG-1999;	9905-0149426.
PR 20-AUG-1999;	9905-0149722.
PR 20-AUG-1999;	9905-0149723.
PR 20-AUG-1999;	9905-0149929.
PR 23-AUG-1999;	9905-0149902.
PR 23-AUG-1999;	9905-0149930.
PR 25-AUG-1999;	9905-0150566.
PR 26-AUG-1999;	9905-0150884.
PR 27-AUG-1999;	9905-0151065.
PR 27-AUG-1999;	9905-0151066.
PR 27-AUG-1999;	9905-0151080.
PR 30-AUG-1999;	9905-0151303.
PR 31-AUG-1999;	9905-0151438.
PR 01-SEP-1999;	9905-0151930.
PR 07-SEP-1999;	9905-0152363.
PR 10-SEP-1999;	9905-0153070.
PR 13-SEP-1999;	9905-0153758.
PR 15-SEP-1999;	9905-0154018.
PR 16-SEP-1999;	9905-0154039.
PR 20-SEP-1999;	9905-0154779.
PR 22-SEP-1999;	9905-0155139.
PR 23-SEP-1999;	9905-0155486.
PR 24-SEP-1999;	9905-0155659.
PR 28-SEP-1999;	9905-0156458.
PR 29-SEP-1999;	9905-0156596.
PR 04-OCT-1999;	9905-0157117.
PR 05-OCT-1999;	9905-0157753.
PR 06-OCT-1999;	9905-0157865.
PR 07-OCT-1999;	9905-0158029.
PR 08-OCT-1999;	9905-0158232.
PR 12-OCT-1999;	9905-0158369.
PR 13-OCT-1999;	9905-0159293.
PR 13-OCT-1999;	9905-0159294.
PR 13-OCT-1999;	9905-0159295.
PR 14-OCT-1999;	9905-0159329.
PR 14-OCT-1999;	9905-0159330.
PR 14-OCT-1999;	9905-0159331.
PR 14-OCT-1999;	9905-0159637.
PR 14-OCT-1999;	9905-0159638.
PR 18-OCT-1999;	9905-0159584.
PR 21-OCT-1999;	9905-0160741.
PR 21-OCT-1999;	9905-0160767.
PR 21-OCT-1999;	9905-0160768.
PR 21-OCT-1999;	9905-0160770.
PR 21-OCT-1999;	9905-0160814.
PR 21-OCT-1999;	9905-0160815.
PR 22-OCT-1999;	9905-0160980.
PR 22-OCT-1999;	9905-0160981.
PR 22-OCT-1999;	9905-0160989.
PR 25-OCT-1999;	9905-0161404.
PR 25-OCT-1999;	9905-0161405.
PR 25-OCT-1999;	9905-0161406.
PR 26-OCT-1999;	9905-0161359.
PR 26-OCT-1999;	9905-0161360.
PR 26-OCT-1999;	9905-0161361.
PR 28-OCT-1999;	9905-0161920.
PR 28-OCT-1999;	9905-0161992.
PR 28-OCT-1999;	9905-0161993.
PR 29-OCT-1999;	9905-0162142.
Query Match	77.1%; Score 37; DB 21; Length 398;

Best Local Similarity	85.7%;	Pred. No. 2,4e+02;							
Matches	6;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	RRRPP	7	:					
Db	161	rrpppp	167						
RESULT	23								
AA606369									
ID	AA606369	standard; Protein;	412	AA.					
XX	AA606369;								
AC									
XX									
DT	17-OCT-2000	(first entry)							
XX									
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 3118.							
XX									
KW	Protein identification; signal transduction pathway; metabolic pathway;								
KW	hybridisation assay; genetic mapping; gene expression control; promoter;								
KW	termination sequence.								
XX									
OS	Arabidopsis thaliana.								
XX	EPI033405-A2.								
PN									
XX									
PD	06-SEP-2000.								
XX									
PF	25-FEB-2000; 2000EP-0301439.								
XX									
PR	25-FEB-1999;	9905-0121825.							

PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139434.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139462.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-0139889.	PR	31-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157733.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	23-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			

PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	31-AUG-1999;	99US-0151303.
PR	01-SEP-1999;	99US-0151438.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157733.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	23-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 77.1%; Score 37; DB 21; Length 412;
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPPP 7
 Db 175 rrrppp 181

XX 04-MAY-1999 (first entry)
 DT Human sucrose isomaltase complex binding protein SM138.
 DE
 XX
 DE Gastro-intestinal transport receptor; binding protein; hsi; hpt1;
 XX D2H; hpept1; human; GI tract receptor; sucrose-isomaltase complex;
 KM intestinal peptide-associated transporter; hypertension; diabetes;
 KM osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KM therapeutic agent delivery; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9651325-A2.
 XX
 PD 19-NOV-1998.
 XX
 XX 15-MAY-1998; 98WO-US10088.
 PF
 XX 15-MAY-1997; 97US-0046595.
 PR
 XX (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin LJ,
 PI Omahony DJ, Patterson CA, Singleton J;
 DR WPI; 1999-009568/01.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 XX through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Claim 2; Page 54; 294pp; English.
 XX
 CC This sequence represents a peptide that specifically binds to the human
 CC sucrose-isomaltase complex. The invention relates to purified
 CC proteins (I) that bind specifically to at least one of the
 CC gastro-intestinal (GI) tract receptors human intestinal
 CC peptide-associated transporter (HPT1), hpept1, D2H and human
 CC sucrose-isomaltase complex (hSI). (I) provide active transport of
 CC therapeutic agents through human and animal GI tissue (into the blood)
 CC for in vivo delivery, particularly for treatment or prevention of
 CC hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,
 CC migraine, or angina pectoris. Specifically they are used to deliver
 CC insulin or leuprolide, but many other suitable therapeutic agents are
 CC disclosed, including genes or inhibitory nucleic acid, imaging agents and
 CC antigens. (I) may also provide targeting to the GI tract. Other uses of
 CC (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 CC
 SQ Sequence 39 AA:

Query Match 75.0%; Score 36; DB 20; Length 39;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRPDP 7
 DB 34 rrpdp 39

RESULT 27
 AAR79134
 ID AAR79134 standard; peptide; 41 AA.
 XX
 AC AAR79134;

XX 27-FEB-1996 (first entry)
 DT W11u cell derived mink kipl, p27-kipl or p27 internal peptide.
 XX
 DE
 XX
 DE Mink kipl; p27-kipl; p27; cyclin E-Cdk2 complex; cancer;
 KM activation inhibitor; hyperplasia; cyclin dependent kinase;
 KM diagnosis; hyperproliferative disorder; ulcer; internal peptide;
 KM W11u cell derived.
 XX
 OS Mustela vison.
 XX
 PN WO9518824-A1.
 XX
 PD 13-JUL-1995.
 XX
 XX 09-JAN-1995; 95WO-US00247.
 PF
 XX 15-JUL-1994; 94US-0275983.
 PR 07-JAN-1994; 94US-0179045.
 PR
 XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 PA (SLOK) SLOAN KETTERING INSTITUTE FOR CANCER RESEARCH.
 XX
 PI Koff A, Massague J, Polyak K, Roberts JM;
 PI WPI; 1995-255037/33.
 DR
 XX
 PT Novel protein, p27 inhibits activation of a cyclin E-Cdk2 complex -
 PT useful for treatment of hyper-proliferative disorders, e.g. cancer,
 PT hyperplasia or ulcers.
 XX
 XX
 PS Disclosure; Page 18; 121pp; English.
 XX
 CC AAR79134 is a W11u cell derived mink kipl (p27-kipl or p27) internal
 CC peptide. p27 inhibits the activation of the cyclin E-Cdk2 (cyclin
 CC dependent kinase) complex. Agents which inhibit or enhance the
 CC ability of p27 to inhibit the activation of cyclin E-Cdk2, can be
 CC used to treat hyperproliferative disorders, e.g. cancer,
 CC hyperplasia or ulcer. Diagnosis of hyperproliferative disorders,
 CC esp. human cancer, can be achieved by detecting a p27 mutation in
 CC the cells of the patient. The disorder can be treated using a
 CC pharmaceutical compsn. comprising a recombinant virus contg. a
 CC nucleic acid mol. encoding p27.
 CC
 SQ Sequence 41 AA:

Query Match 75.0%; Score 36; DB 16; Length 41;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRPDP 8
 DB 30 rrpdp 36

RESULT 28
 AAR50053
 ID AAR50053 standard; Protein; 46 AA.
 XX
 AC AAR50053;
 XX
 XX 04-MAY-1994 (first entry)
 DT ICP34.5 fragment.
 XX
 DE
 XX
 XX Gamma1 34.5; ICP34.5; programmed cell death; apoptosis;
 KM neurovirulence; inhibition; viral replication; gene therapy;
 KM neurodegenerative disease; herpes simplex virus; ss.
 XX
 OS Herpes simplex virus 1.
 XX

FN W09319591-A.
XX 14-OCT-1993.
XX 26-FEB-1993; 93WO-US01801.
XX 31-MAR-1992; 92US-0861233.
XX (ARCH-) ARCH DEV CORP.
XX Chou J, Roizman B;
XX WPI: 1993-336453/42.
XX Use of gamma, 34.5 gene or encoded ICP 34.5 - for preventing or
PT treating programmed cell death in neuronal cells or in screening
PT assays
PS Disclosure: Page 72; 95pp; English.
XX Programmed cell death may be treated through the use of the HSV-1
CC gene gamma1 34.5 or prod. of its expression, ICP34.5. The gene and
CC its expression have been demonstrated to be required for HSV-1
CC neurovirulence, and in partic., to act as an inhibitor of neuronal
CC programmed cell death which allows for viral replication. Use of
CC the gene therapy, or the protein itself, can be expected to result
CC in inhibition of programmed cell death in various neurodegenerative
CC diseases.
CC DNA sequences of HSV-1 strains F (AA050037-38, AA050040-42),
CC 17syn+ (AA050043-45, AA058777), MGH-10 (AA058778-83), and CVG-2
CC (AA058784-88) and the predicted open reading frames for ICP34.5
CC (AA050048-58) are compared in Fig 1.
XX Sequence 46 AA;

Query Match 75.0%; Score 36; DB 14; Length 46;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
|||||
Db 12 rrrpp 17

RESULT 29

AA010803 standard; Protein; 70 AA.

XX AAG10803;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9270.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 98US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137322.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

CC The present sequence is the protein produced from the murine prostate
CC cancer predisposing gene HPC2, the human homologue of which is found on
CC chromosome 17p. Some alleles of this gene cause a predisposition to
CC cancer, particularly prostate cancer. This protein and its gene can be
CC used in peptide and gene therapy for cancer patients, as well as being
CC useful as diagnostic tools (both for cancer sufferers and those with a
CC predisposition to the disease) and in the production of cancer drugs.

XX Sequence 73 AA;

Query Match 75.0%; Score 36; DB 21; Length 73;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RRRPP 7
 |||||
Db 22 rrrpp 27

RESULT 31

AA010802
ID AAG10802 standard; Protein; 76 AA.

XX AAG10802;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9269.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9269.
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

XX EPI03405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136382.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145182.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR	02-AUG-1999	9905-0147038
PR	03-AUG-1999	9905-0147089
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149378
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149723
PR	27-AUG-1999	9905-0149929
PR	28-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150566
PR	26-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0153263
PR	10-SEP-1999	9905-0153707
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155134
PR	23-SEP-1999	9905-0155186
PR	24-SEP-1999	9905-0155559
PR	28-SEP-1999	9905-0155658
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158365
PR	13-OCT-1999	9905-0158293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159339
PR	14-OCT-1999	9905-0159631
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160815
PR	21-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	25-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161920

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match	75.0%	Score 36	DB 21	length 76
Best Local Similarity	85.7%	Pred. No. 79		
Matches	6	Conservative	0	Mismatches 1; Indels 0; Gaps 0;

QY	2	RRPRPPY	8
Db	70	rrpprpy	76

RESULT 32

AAB25254 standard; Protein; 155 AA.

AC AAB25254

DT 27-NOV-2000 (first entry)

DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:573.

KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signaling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism.

OS Eucalyptus grandis.

PN WO200042171-A1.

PD 20-JUL-2000.

PF 11-JAN-2000; 2000WO-US00724.

PR 12-JAN-1999; 99US-0228986.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strabala TJ, Nieuwenhuizen NJ;

DR WPI; 2000-476052/41.

PT Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals -

PS Claim 3; Page 257; 527pp; English.

CC AAA79266 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or
CC pine (*Pinus radiata* also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.

50 Sequence 155 AA;

Query Match	75.08;	Score 36;	DB 21;	Length 155;
-------------	--------	-----------	--------	-------------

Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
|||||
DB 127 rrrpp 132

RESULT 33

AAG26907 standard; Protein; 166 AA.

AC AAG26907;

DT 17-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 31542.

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.

OS Zea mays subsp. mays.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130049.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135533.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 75.0%; Score 36; DB 21; Length 166;
Best Local Similarity 100.0%; Fred.No. 1; 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
DB 47 rrrpp 52

RESULT 34
AAG27133 standard; Protein: 166 AA.
ID AAG27133
AC AAG27133;
XX 17-OCT-2000 (first entry)
DE Zea mays protein fragment SEQ ID NO: 31852.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147600.
PR 06-AUG-1999; 99US-0147703.
PR 06-AUG-1999; 99US-0147716.
PR 09-AUG-1999; 99US-0147793.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 75.0%; Score 36; DB 21; Length 166;
Best local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 125 rrrpp 130

RESULT 35
AAR50051
ID AAR50051 standard; Protein; 169 AA.
XX
AC AAR50051;
XX

DT 04-MAY-1994 (first entry)
 XX ICP34.5 fragment.
 DE
 XX Gammal 34.5; ICP34.5; programmed cell death; apoptosis;
 KW neurovirulence; inhibition; viral replication; gene therapy;
 XX neurodegenerative disease; herpes simplex virus; ss.
 XX Herpes simplex virus 1.
 OS
 XX WO9319591-A.
 PN
 XX 14-OCT-1993.
 PD
 XX 26-FEB-1993; 93WO-US01801.
 PF
 XX 31-MAR-1992; 92US-0861233.
 PR
 XX (ARCH-) ARCH DEV CORP.
 PA
 XX Chou J, Rolzman B;
 PI
 XX WPI; 1993-336453/42.
 DR
 XX Use of gamma, 34.5 gene or encoded ICP 34.5 - for preventing or
 PT treating programmed cell death in neuronal cells or in screening
 PT assays
 PS
 XX Disclosure: Page 70-71; 95pp; English.
 CC Programmed cell death may be treated through the use of the HSV-1
 CC gene gammal 34.5 or prod. of its expression, ICP34.5. The gene and
 CC its expression have been demonstrated to be required for HSV-1
 CC neurovirulence, and in partic., to act as an inhibitor of neuronal
 CC programmed cell death which allows for viral replication. Use of
 CC the gene therapy, or the protein itself, can be expected to result
 CC in inhibition of programmed cell death in various neurodegenerative
 CC diseases.
 CC DNA sequences of HSV-1 strains F (AAO50037-38, AAO50040-42),
 CC 17syn+ (AAO50043-45, AAO58777), MGH-10 (AAO58778-83), and CVG-2
 CC (AAO58784-88) and the predicted open reading frames for ICP34.5
 CC (AAO50048-58) are compared in Fig 1.
 CC
 XX
 SQ Sequence 169 AA:
 QY
 2 RRRPP 7
 |||||
 Db 5 rrrpp 10
 Query Match 75.0%; Score 36; DB 14; Length 169;
 Best Local Similarity 100.0%; Pred. NO. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PD 14-OCT-1993.
 XX
 PF 26-FEB-1993; 93WO-US01801.
 XX
 PR 31-MAR-1992; 92US-0861233.
 XX
 PA (ARCH-) ARCH DEV CORP.
 XX
 XX Chou J, Rolzman B;
 PI
 XX WPI; 1993-336453/42.
 DR
 XX Use of gamma, 34.5 gene or encoded ICP 34.5 - for preventing or
 PT treating programmed cell death in neuronal cells or in screening
 PT assays
 PS
 XX Disclosure: Page 74-75; 95pp; English.
 CC Programmed cell death may be treated through the use of the HSV-1
 CC gene gammal 34.5 or prod. of its expression, ICP34.5. The gene and
 CC its expression have been demonstrated to be required for HSV-1
 CC neurovirulence, and in partic., to act as an inhibitor of neuronal
 CC programmed cell death which allows for viral replication. Use of
 CC the gene therapy, or the protein itself, can be expected to result
 CC in inhibition of programmed cell death in various neurodegenerative
 CC diseases.
 CC DNA sequences of HSV-1 strains F (AAO50037-38, AAO50040-42),
 CC 17syn+ (AAO50043-45, AAO58777), MGH-10 (AAO58778-83), and CVG-2
 CC (AAO58784-88) and the predicted open reading frames for ICP34.5
 CC (AAO50048-58) are compared in Fig 1.
 CC
 XX
 SQ Sequence 179 AA:
 QY
 2 RRRPP 7
 |||||
 Db 12 rrrpp 17
 Query Match 75.0%; Score 36; DB 14; Length 179;
 Best Local Similarity 100.0%; Pred. NO. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 37
 ID AAY73974 standard; Protein; 235 AA.
 XX
 AC AAY73974;
 XX
 DT 14-MAR-2000 (first entry)
 DE
 XX Human prostate tumor EST fragment derived protein #161.
 XX
 KW Pancreas; tumor; EST; expressed sequence tag; human; cytosstatic;
 KW treatment.
 OS
 XX Homo sapiens.
 XX
 PN DE19820190-A1.
 XX
 XX 04-NOV-1999.
 PD
 XX 28-APR-1998; 98DE-1020190.
 PF
 XX 28-APR-1998; 98DE-1020190.
 PR
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
 PI
 XX WPI; 1999-621386/54.
 DR
 XX N-PSDB; AA252911.
 XX


```

RESULT 40
OS AAR21706
ID AAR21706 standard; Protein; 263 AA.
XX
AC AAR21706;
XX
DT 28-JUL-1992 (first entry)
XX
DE HSV-1 (F) ICP34.5 protein.
XX
KW Herpes simplex virus; avirulent; vaccine; immunisation.
XX
OS Herpes simplex virus.
XX
PN WO9204050-A.
XX
PD 19-MAR-1992.
XX
PF 10-SEP-1991; 91WO-US06532.
XX
PR 10-SEP-1990; 90US-0579834.
XX
PA (ROIZ/) ROIZMAN B.
XX
PI Roizman B;
XX
DR WPI; 1992-114074/14.
DR N-PSDB; AAQ23295.
XX
PT New recombinant Herpes Simplex Virus vaccines - rendered
XX avirulent by deletion of ICP 34.5 gene encoding active gene prod.
XX
PS Disclosure; Page 44; 78pp; English.
XX
CC Herpes simplex virus-1 (F) was isolated from a recurrent facial
CC vesicle and passaged a maximum of four times in cells in culture.
CC Viral DNA was prep'd from virions that accumulated in the cytoplasm
CC of infected VERO cells (Kieff et al., J. Virol., 8, 125-132 (1971)).
CC The BamHI SP junction fragments contg. the domain of the gene that
CC specified ICP34.5 were cloned into a pUC18 plasmid using HSV-1(F)
CC sequences as probes in colony blot hybridisation. Further
CC subclonings were done to facilitate sequencing of different regions
CC of the gene. HSV may be rendered avirulent by removing the ICP34.5
CC gene from the viral genome. The modified HSV genome is useful for
CC preparing a stable, non-transforming live viral vaccine which
CC either does not establish latent infections or which cannot be
CC reactivated from a latent state.
CC See also AAR21707,8.
XX
SQ Sequence 263 AA:

```

Query Match 75.0%; Score 36; DB 13; Length 263;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 RRRPP 7
   |||||
DB 11 rrrpp 16

```

RESULT 41
 AAR78729 standard; Protein; 294 AA.
 XX
 AC AAR78729;
 XX
 DT 22-NOV-1995 (first entry)
 XX
 DE Human bone morphogenetic protein (BMP)-12.
 XX
 KW Bone morphogenetic protein; BMP-12; tendon; ligament; tendinitis; ss.

```

XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT 1..190
XX FT /label= propeptide
XX FT 191..294
XX FT Protein
XX FT /label= mature sequence
XX
PN WO9516035-A.
XX
PD 15-JUN-1995.
XX
PF 06-DEC-1994; 94WO-US14030.
XX
PR 02-NOV-1994; 94US-0333576.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
XX
PA (GEM ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX
DR WPI; 1995-224320/29.
DR N-PSDB; AAQ96207.
XX
PT Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in
PT compsn. for inducing tendon/ligament-like tissue formation
XX
PS Claim 14; Page 50-51; 84pp; English.
XX
CC Oligo probe 3 (AAQ96214) designed on the basis of an amplified BMP-12
CC human DNA sequence (AAQ96212) was radioactively labelled and used to
CC screen a human genomic library. A positive clone was designated
CC lambda-Huc-48. This bacteriophage was deposited under ATCC 75625 on 7
CC Dec. 1993. The oligo hybridising region of this phage is localised
CC to a 3.2 kb BamHI fragment which was subcloned to PCR1-1#2 and
CC deposited under ATCC 69517 on 7 Dec. 1993. The partial DNA sequence
CC and derived AA sequence of the 3.2 kb DNA insert of the plasmid
CC subclone PCR1-1#2, derived from clone lambda-Huc-48 are shown in
CC AAQ96207 and AAR78729 respectively. It includes the entire mature
CC sequence and approx. 190 AAs of the propeptide. The CDS of mature
CC BMP-12 begins at nt. #496 or #571. The first Cys in the seven Cys
CC structure characteristic of TGF-beta proteins begins at nt. #577.
CC The last Cys ends at #879. Thus it is expected that DNA sequences
CC encoding active BMP-12 species will comprise nts #577 to #879 (see
CC also AAQ96207 FT).
XX
SQ Sequence 294 AA:

```

Query Match 75.0%; Score 36; DB 16; Length 294;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 RRRPP 7
   |||||
DB 65 rrrpp 71

```

RESULT 42
 AAW26589 standard; Protein; 294 AA.
 XX
 AC AAW26589;
 XX
 DT 21-JAN-1998 (first entry)
 XX
 DE Human bone morphogenetic protein BMP-12.
 XX
 KW BMP-12; bone morphogenetic protein; human; tendon; ligament;

[illegible]

FT		/label= Mat_protein
XX		
PN	US5658882-A.	
XX		
PD	19-AUG-1997.	
XX		
PF	07-DEC-1993;	93US-0164103.
XX		
PR	22-DEC-1994;	94US-0362670.
PR	07-DEC-1993;	93US-0164103.
PR	25-MAR-1994;	94US-0217780.
PR	02-NOV-1994;	94US-0333576.
XX		
PA	(GENEY) GENETICS INST INC.	
PA	(HARD) HARVARD COLLEGE.	
XX		
PI	Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;	
PI	Wozney JM;	
XX		
DR	WPI; 1997-424270/39.	
DR	N-PSDB; AAT90390.	
XX		
PT	- Inducing tendon and ligament formation using BMP-12, BMP-13 or MP-52	
PT	- useful for tissue healing and repair, treatment of tendonitis,	
XX	improving fixation of tendons to bone etc	
XX		
PS	Example 1; Column 77-80; 43pp; English.	
XX		
CC	This polypeptide comprises a novel bone morphogenetic protein,	
CC	designated BMP-12, that induces tendon and ligament formation. Its	
CC	amino acid sequence was deduced from isolated genomic clone VI-1	
CC	(see AAT90390). A claimed method for inducing formation of tendon	
CC	and/or ligament tissues involves the administration of a	
CC	composition containing at least one protein selected from BMP-12	
CC	(see also AAW26589), MP52 (see AAW26590) and BMP-13 (see AAW26591).	The
CC	Method is useful for tissue (including skin) healing and repair.	
CC	This is useful for treating tendonitis, carpal tunnel syndrome and	
CC	other defects of traumatic or congenital origin, in cosmetic	
CC	surgery and to improve fixation of tendons or ligaments to bone.	
CC	The specified protein/enzyme also be used to increase activity of other	
XX	BMPs e.g. BMP-2 (see AAW26597).	
XX		
SQ	Sequence 388 AA;	
OY	Query Match 75.0%; Score 36; DB 18; Length 388;	
	Best Local Similarity 85.7%; Pred. No. 3.3e+02;	
Db	Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
	159 rreptp 165	

Search completed: September 24, 2001, 10:08:37
Job time: 197 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 24, 2001, 10:06:41 : Search time 26.42 Seconds
(without alignments)
6.235 Million cell updates/sec

Title: US-09-276-868-5

Perfect score: 48

Sequence: 1 RRRPRPPY 8

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	14	4	US-09-024-975-4
2	48	100.0	26	2	US-08-419-066-2
3	48	100.0	26	4	US-09-024-975-2
4	48	100.0	39	1	US-08-162-052-1
5	48	100.0	39	1	US-08-310-722-1
6	48	100.0	39	2	US-08-419-066-1
7	48	100.0	39	2	US-08-728-333-1
8	48	100.0	39	4	US-09-024-975-1
9	48	100.0	39	5	PCR-US95-12080-1
10	36	75.0	29	3	US-08-256-747C-45
11	36	75.0	29	4	US-08-834-130A-45
12	36	75.0	46	4	US-08-483-533-30
13	36	75.0	169	4	US-08-483-533-28
14	36	75.0	179	4	US-08-483-533-33
15	36	75.0	245	4	US-08-483-533-42
16	36	75.0	252	4	US-08-483-533-43
17	36	75.0	258	4	US-08-483-533-26
18	36	75.0	263	5	PCR-US91-06532-2
19	36	75.0	264	4	US-08-483-533-40
20	36	75.0	277	6	5164481-2
21	36	75.0	294	1	US-08-362-670B-2
22	36	75.0	294	3	US-08-333-576C-2
23	36	75.0	294	5	PCR-US94-14030A-2
24	36	75.0	355	4	US-08-483-533-41
25	36	75.0	355	5	PCR-US91-06532-3
26	36	75.0	388	1	US-08-362-670B-34
27	36	75.0	388	3	US-08-333-576C-34

28	36	75.0	388	5	PCR-US94-14030A-34	Sequence 34, Appl
29	36	75.0	992	1	US-08-127-499A-1	Sequence 1, Appl
30	36	75.0	992	1	US-08-482-847-1	Sequence 1, Appl
31	36	75.0	1063	1	US-08-093-453B-3	Sequence 3, Appl
32	36	75.0	1063	1	US-08-127-499A-8	Sequence 8, Appl
33	36	75.0	1063	1	US-08-482-847-8	Sequence 8, Appl
34	34	70.8	180	3	US-09-187-331-5	Sequence 5, Appl
35	34	70.8	195	3	US-09-187-331-1	Sequence 1, Appl
36	34	70.8	326	1	US-08-293-563-7	Sequence 7, Appl
37	34	70.8	454	3	US-08-434-099A-27	Sequence 27, Appl
38	34	70.8	454	3	US-08-434-099A-28	Sequence 28, Appl
39	34	70.8	454	4	US-08-764-870-5	Sequence 5, Appl
40	34	70.8	454	6	5260432-2	Patent No. 5260432
41	34	70.8	520	2	US-09-091-432-2	Sequence 2, Appl
42	34	70.8	2293	4	US-09-368-590-2	Sequence 2, Appl
43	33.5	69.8	280	4	US-09-264-419C-2	Sequence 2, Appl
44	33	68.8	20	4	US-09-024-975-9	Sequence 9, Appl
45	33	68.8	59	5	PCR-US95-12080-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-024-975-4

Sequence 4, Application US/09024975

Patent No. 6133233

GENERAL INFORMATION:

APPLICANT: ROSS, CHRISTOPHER R.

APPLICANT: BLECHA, FRANK

APPLICANT: SHI, JISHU

TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS

STREET: 2405 GRAND BLVD., SUITE 400

CITY: KANSAS CITY

STATE: MO

COUNTRY: USA

ZIP: 64108

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/024, 975

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/802, 306

FILING DATE: 18-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: COLLINS, JOHN M.

REGISTRATION NUMBER: 26,262

REFERENCE/DOCKET NUMBER: 25585-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 816/474-9050

TELEFAX: 816/474-9057

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-024-975-4

Query Match 100.0%; Score 48; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPPY 8

Db 1 RRRPRPY 8

RESULT 2

US-08-419-066-2
; Sequence 2, Application US/08419066
; Patent No. 5830993
; GENERAL INFORMATION:
; APPLICANT: Blecha, Frank
; APPLICANT: Shi, Jishu
; TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John M. Collins, Hovey, Williams, Timmons &
; ADDRESSEE: Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,066
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26262
; REFERENCE/DOCKET NUMBER: 23625
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (816) 474-9050
; TELEFAX: (816) 474-9057
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-419-066-2

Query Match 100.0%; Score 48; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPY 8
Db 1 RRRPRPY 8

RESULT 3

US-09-024-975-2
; Sequence 2, Application US/09024975
; Patent No. 6133233
; GENERAL INFORMATION:
; APPLICANT: ROSS, CHRISTOPHER R.
; APPLICANT: BLECHA, FRANK
; APPLICANT: SHI, JISHU
; TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
; STREET: 2405 GRAND BLVD., SUITE 400
; CITY: KANSAS CITY

STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-2

Query Match 100.0%; Score 48; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPY 8
Db 1 RRRPRPY 8

RESULT 4
US-08-162-052-1
; Sequence 1, Application US/08162052
; Patent No. 5489575
; GENERAL INFORMATION:
; APPLICANT: LEE, Jong-Youn
; APPLICANT: BOMAN, Hans G
; APPLICANT: MUTT, Viktor
; APPLICANT: JORNVAL, Hans
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND THEIR USE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22133-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,052
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9101838-2
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92-22578
; FILING DATE: 23-DEC-1992

ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 003300-299
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-162-052-1

Query Match 100.0%; Score 48; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
DB 1 RRRRPPY 8

RESULT 5
US-08-310-722-1
; Sequence 1, Application US/08310722
; Patent No. 5654273
; GENERAL INFORMATION:
; APPLICANT: Gallo, Richard L.
; APPLICANT: Klagsbrun, Michael
; TITLE OF INVENTION: Syncytin Mediated Modulation of Tissue Repair
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/310,722
; FILING DATE: 22-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: CMC379
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PUBLICATION INFORMATION:
; AUTHORS: Lee, Jong-Youn
; AUTHORS: Boman, Hans G.
; AUTHORS: Mutt, Viktor
; AUTHORS: Jorntall, Hans
; TITLE: NO. 5654273el Polypeptides And Their Use
; JOURNAL: PCT WO 92/22578

DATE: 12/23/92
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
; US-08-310-722-1

Query Match 100.0%; Score 48; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
DB 1 RRRRPPY 8

RESULT 6
US-08-419-066-1
; Sequence 1, Application US/08419066
; Patent No. 5830993
; GENERAL INFORMATION:
; APPLICANT: Blecha, Frank
; APPLICANT: Shi, Jishu
; TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John M. Collins, Hovey, Williams, Timmons &
; ADDRESSEE: Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,066
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26262
; REFERENCE/DOCKET NUMBER: 23625
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (816) 474-9050
; TELEFAX: (816) 474-9057
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-419-066-1

Query Match 100.0%; Score 48; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
DB 1 RRRRPPY 8

RESULT 7
US-08-728-333-1
; Sequence 1, Application US/08728333
; Patent No. 5863897

GENERAL INFORMATION:
APPLICANT: Gallo, Richard L.
APPLICANT: Klagsbrun, Michael
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,333
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/310,722
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: CMCC379
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Mult, Viktor
AUTHORS: Jorinval, Hans
TITLE: No. 5863897el Polypeptides And Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
US-08-728-333-1

Query Match 100.0%; Score 48; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRPRPPY 8
Db 1 RRRPRPPY 8

RESULT 8
US-09-024-975-1
Sequence 1, Application US/09024975
Patent No. 6133233
GENERAL INFORMATION:
APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400

CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-1

Query Match 100.0%; Score 48; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRPRPPY 8
Db 1 RRRPRPPY 8

RESULT 9
PCT-US95-12080-1
Sequence 1, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Lee, Jong-Youn
AUTHORS: Boman, Hans G.
AUTHORS: Milt, Viktor
AUTHORS: Jorvall, Hans
TITLE: Novel Polypeptides And Their Use
JOURNAL: PCT WO 92/22578
DATE: 12/23/92
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 39
PCT-US95-12080-1

Query Match 100.0%; Score 48; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 RRRPP 8
DB 1 RRRPP 8

RESULT 10
US-08-256-747C-45
Sequence 45, Application US/08256747C
Patent No. 6037448
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: GILLAM, Shirley
APPLICANT: OU, Dawei
APPLICANT: TINGLE, Aubrey
TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,747C
FILING DATE: 06-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-370
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-747C-45

Query Match 75.0%; Score 36; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 RRRPP 7
DB 1 RRRPP 7

DB 8 RRRPP 13

RESULT 11
US-08-834-130A-45
Sequence 45, Application US/08834130A
Patent No. 6180758
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: GILLAM, Shirley
APPLICANT: OU, Dawei
APPLICANT: TINGLE, Aubrey
TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,130A
FILING DATE: 14-APR-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-686 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-834-130A-45

Query Match 75.0%; Score 36; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 RRRPP 7
DB 8 RRRPP 13

RESULT 12
US-08-483-533-30
Sequence 30, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Rolzman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
DISEASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-30

Query Match 75.0%; Score 36; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 12 RRRPP 17

RESULT 13
US-08-483-533-28
Sequence 28, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
TITLE OF INVENTION: Diseases
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92

ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 169 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-28

Query Match 75.0%; Score 36; DB 4; Length 169;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 5 RRRPP 10

RESULT 14
US-08-483-533-33
Sequence 33, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
TITLE OF INVENTION: Diseases
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 179 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-483-533-33

Query Match 75.0%; Score 36; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 12 RRRPP 17

RESULT 15

US-08-483-533-42
; Sequence 42, Application US/08483533
; Patent No. 6172047
; GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,533
; FILING DATE: 07-MAR-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/419,853
; FILING DATE: 11-APR-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/861,233
; FILING DATE: 31-MAR-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, James P.
; REGISTRATION NUMBER: 28,491
; REFERENCE/DOCKET NUMBER: 28097/32742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-533-42

Query Match 75.0%; Score 36; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 12 RRRPP 17

RESULT 16

US-08-483-533-43
; Sequence 43, Application US/08483533
; Patent No. 6172047

GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-43

Query Match 75.0%; Score 36; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
Db 12 RRRPP 17

RESULT 17

US-08-483-533-26
; Sequence 26, Application US/08483533
; Patent No. 6172047

GENERAL INFORMATION:
; APPLICANT: Roizman, Bernard
; APPLICANT: Chou, Joany
; TITLE OF INVENTION: Method for Treating Tumorigenic
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois

Query Match 75.0%; Score 36; DB 4; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 258 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-26

Query Match 75.0%; Score 36; DB 4; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
111111
Db 5 RRRPP 10

RESULT 18
PCT-US91-06532-2
Sequence 2, Application PC/TUS9106532
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
TITLE OF INVENTION: Vaccines and Methods
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Bicknell
STREET: Two First National Plaza Suite 2100
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06532
FILING DATE: 19910910
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gruber, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27373/8235

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-06532-2

Query Match 75.0%; Score 36; DB 5; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
111111
Db 11 RRRPP 16

RESULT 19
US-08-483-533-40
Sequence 40, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
APPLICANT: Chou, Joany
TITLE OF INVENTION: Method for Treating Tumorigenic
TITLE OF INVENTION: Diseases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-40

Query Match 75.0%; Score 36; DB 4; Length 264;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
Db 11 RRRPP 16

RESULT 20
5164481-2
; Patent No. 5164481
; APPLICANT: LACROIX, MARTIAL; REIN, MAAN; DIONNE, GERVAIN
; TITLE OF INVENTION: PEPTIDES AND ANALOGUES AND MIXTURES
; FOR DETECTING AND ELICITING ANTIBODIES TO RUBELLA VIRUS
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/397,767
; FILING DATE: 23-AUG-1989
; SEQ ID NO: 2:
; LENGTH: 277
5164481-2

Query Match 75.0%; Score 36; DB 6; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
Db 35 RRRPP 40

RESULT 21
US-08-362-670B-2
; Sequence 2, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Mooney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolman, Neil
; APPLICANT: Thomson, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-362-670B-2

Query Match 75.0%; Score 36; DB 1; Length 294;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPP 7
Db 65 RRRPP 71

RESULT 22
US-08-333-576C-2
; Sequence 2, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Mooney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolman, Neil
; APPLICANT: Thomson, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919 September 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ. ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-576C-2

Query Match 75.0%; Score 36; DB 3; Length 294;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPP 7
Db 65 RRRPP 71

RESULT 23
PCT-US94-14030A-2
; Sequence 2, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE

TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14030A
FILING DATE: Herewith
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/164,103
FILING DATE: 07-DEC-1993
APPLICATION NUMBER: US 08/217,780
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: US 08/333,576
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202D-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 294 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-14030A-2

Query Match 75.0%; Score 36; DB 5; Length 294;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRPP 7
DB 65 RRRRPP 71

RESULT 24
US-08-483-533-41
Sequence 41, Application US/08483533
Patent No. 6172047
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Method for Treating Tumorigenic
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,533
FILING DATE: 07-MAR-95
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/419,853
FILING DATE: 11-APR-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/861,233
FILING DATE: 31-MAR-92
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, James P.
REGISTRATION NUMBER: 28,491
REFERENCE/DOCKET NUMBER: 28097/32742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-533-41

Query Match 75.0%; Score 36; DB 4; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRRPP 7
DB 11 RRRRPP 16

RESULT 25
PCT-US91-06532-3
Sequence 3, Application PC/TUS9106532
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blacknell, O'Toole, Gerstein, Murray &
STREET: Two First National Plaza Suite 2100
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06532
FILING DATE: 19910910
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gruber, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27373/8235
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: AMINO ACID

TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-06532-3

Query Match 75.0%; Score 36; DB 5; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 RRRPP 7
111111
DB 11 RRRPP 16

RESULT 26
US-08-362-670B-34
Sequence 34, Application US/08362670B
Patent No. 5658882
GENERAL INFORMATION:

APPLICANT: Celeste, Anthony J.
APPLICANT: Mooney, John
APPLICANT: Rosen, Vicki A.
APPLICANT: Wolman, Neil
APPLICANT: Thomsen, Gerald H.
APPLICANT: Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,670B
FILING DATE: December 22, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-670B-34

Query Match 75.0%; Score 36; DB 1; Length 388;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
111111
DB 159 RRRPP 165

RESULT 27
US-08-333-576C-34
Sequence 34, Application US/08333576C
Patent No. 6027919
GENERAL INFORMATION:

APPLICANT: Celeste, Anthony J.
APPLICANT: Mooney, John
APPLICANT: Rosen, Vicki A.
APPLICANT: Wolman, Neil
APPLICANT: Thomsen, Gerald H.
APPLICANT: Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESS: GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/333,576C
FILING DATE: No. 6027919eember 2, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-333-576C-34

Query Match 75.0%; Score 36; DB 3; Length 388;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
111111
DB 159 RRRPP 165

RESULT 28
PCT-US94-14030A-34
Sequence 34, Application PC/TUS9414030A
GENERAL INFORMATION:

APPLICANT: GENETICS INSTITUTE, INC.
APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14030A
FILING DATE: Herewith
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/164,103
FILING DATE: 07-DEC-1993
APPLICATION NUMBER: US 08/217,780
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: US 08/333,576
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202D-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 876-5851
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-14030A-34

Query Match 75.0%; Score 36; DB 5; Length 388;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRRPP 7
Db 159 RRRRPP 165

RESULT 29
US-08-127-499A-1
Sequence 1, Application US/08127499A
Patent No. 5510264
GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
NUMBER OF INVENTION: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,499A
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/102/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-127-499A-1

Query Match 75.0%; Score 36; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRRPP 7
Db 35 RRRRPP 40

RESULT 30
US-08-482-847-1
Sequence 1, Application US/08482847
Patent No. 5556757
GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,847
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/127,499
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/104/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 992 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-482-847-1

Query Match 75.0%; Score 36; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRRPP 7
Db 35 RRRRPP 40

RESULT 31
US-08-093-453B-3
Sequence 3, Application US/08093453B
Patent No. 5439814
GENERAL INFORMATION:
APPLICANT: Frey, Terry K.
APPLICANT: Dominguez, Geraldina

APPLICANT: Wang, Chin Yen
TITLE OF INVENTION: Modified Infectious Rubella Virus
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jamie L. Greene, Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: United States
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,453B
FILING DATE: 19 JUL 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 07/722,334
FILING DATE: 28 JUN 1991
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 07362-0101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404_818-3700
TELEFAX: 404_818-3799
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE: Rubella virus
ORGANISM: Thierien
US-08-093-453B-3

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RRRPP 7
Db 35 RRRPP 40

RESULT 32
US-08-127-499A-8
Sequence 8, Application US/08127499A
Patent No. 5510264
GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,499A
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/102/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-127-499A-8

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RRRPP 7
Db 35 RRRPP 40

RESULT 33
US-08-482-847-8
Sequence 8, Application US/08482847
Patent No. 5556757
GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,847
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/127,499
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/104/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-482-847-8

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
11111
DB 35 RRRPP 40

RESULT 34
US-09-187-331-5
Sequence 5, Application US/09187331
Patent No. 6043056
GENERAL INFORMATION:

APPLICANT: Yue, Henry
APPLICANT: Cortley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
FILE REFERENCE: PF-0631 US

CURRENT APPLICATION NUMBER: US/09/187,331
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6

SOFTWARE: PERL Program
SEQ ID NO 5

LENGTH: 180

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE: -
OTHER INFORMATION: g2499136

US-09-187-331-5

Query Match 70.8%; Score 34; DB 3; Length 180;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
11111
DB 100 RRRPP 106

RESULT 35

US-09-187-331-1
Sequence 1, Application US/09187331
Patent No. 6043056

GENERAL INFORMATION:

APPLICANT: Yue, Henry
APPLICANT: Cortley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Gorgone, Gina A.

APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS

FILE REFERENCE: PF-0631 US
CURRENT APPLICATION NUMBER: US/09/187,331

CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 6

SOFTWARE: PERL Program
SEQ ID NO 1

LENGTH: 195

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE: -
OTHER INFORMATION: 2297891

US-09-187-331-1

Query Match 70.8%; Score 34; DB 3; Length 195;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
11111
DB 100 RRRPP 106

RESULT 36
US-08-293-563-7
Sequence 7, Application US/08293563
Patent No. 5516894
GENERAL INFORMATION:

APPLICANT: Reppert, Steven M.
TITLE OF INVENTION: A2b-ADENOSINE RECEPTORS AND
RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: Wordperfect (Version 5.1)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,563

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/850,188

FILING DATE: March 11, 1992

ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/128001

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 326

TYPE: amino acid

STRANDEDNESS: N/A

TOPOLOGY: linear

US-08-293-563-7

Query Match 70.8%; Score 34; DB 1; Length 326;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPP 7
11111
DB 308 RRRPP 314

RESULT 37
US-08-434-099A-27
Sequence 27, Application US/08434099A
Patent No. 6083902

GENERAL INFORMATION:
APPLICANT: Cedetholm-Tms., Stewart A.

TITLE OF INVENTION: Recombinant Fibrin Chains,
Fibrin and Fibrin-Homologs

NUMBER OF SEQUENCES: 37

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: E.R. Squibb & Sons, Inc.
;; STREET: 100 Headquarters Park Drive
;; CITY: Skillman
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08558
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FASTSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/434,099A
;; FILING DATE: 03-MAY-1995
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/236,979
;; FILING DATE: 02-MAY-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Furman, Jr., Esq., Theodore R
;; REGISTRATION NUMBER: 30,942
;; REFERENCE/DOCKET NUMBER: CV0054a
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-281-2372
;; TELEFAX: 908-281-2373
;;
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 27:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 454 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: Internal
;;
US-08-434-099A-27

Query Match 70.8%; Score 34; DB 3; Length 454;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPRP 7
| | | | |
Db 6 RGRPRP 12

RESULT 38
US-08-434-099A-28
; Sequence 28, Application US/08434099A
; Patent No. 6083902
; GENERAL INFORMATION:
; APPLICANT: Cederholm-Wms., Stewart A.
; TITLE OF INVENTION: Recombinant Fibrin Chains,
; TITLE OF INVENTION: Fibrin and Fibrin-Homologs
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.R. Squibb & Sons, Inc.
; STREET: 100 Headquarters Park Drive
; CITY: Skillman
; STATE: NJ
; COUNTRY: USA
; ZIP: 08558
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,099A
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 08/236,979
;; FILING DATE: 02-MAY-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Furman, Jr., Esq., Theodore R
;; REGISTRATION NUMBER: 30,942
;; REFERENCE/DOCKET NUMBER: CV0054a
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-281-2372
;; TELEFAX: 908-281-2373
;;
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 454 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
US-08-434-099A-28

Query Match 70.8%; Score 34; DB 3; Length 454;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPRP 7
| | | | |
Db 6 RGRPRP 12

RESULT 39
US-08-764-870-5
; Sequence 5, Application US/08764870
; Patent No. 6236946
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Thomas S
; APPLICANT: Baxter, John D
; APPLICANT: Fletcher, Robert J
; APPLICANT: Wagner, Richard L
; APPLICANT: Kushner, Peter J
; APPLICANT: Apriletti, James W
; APPLICANT: West, Brian
; TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
; TITLE OF INVENTION: Binding Domains
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,870
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,540
; FILING DATE: 13-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,543
; FILING DATE: 13-DEC-1995
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,606
; FILING DATE: 14-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Jackie N
; REGISTRATION NUMBER: 35,966
; REFERENCE/DOCKET NUMBER: UCAL-246/01US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (650)843-5000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-764-870-5

Query Match 70.8%; Score 34; DB 4; Length 454;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
||| |
DB 367 RRRPSQPY 374

RESULT 40
5260432-2
; Patent No. 5260432
; APPLICANT: TAKAKU, FUMIMARO, ISHIKAWA, TAKASHI, IMAMARI, MICHIO;
; EVANS, RONALD M.; OMESONO, KAZUHIKO
; TITLE OF INVENTION: HUMAN GAMMA RETINOIC ACID RECEPTOR DNA
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/370,407
; FILING DATE: 22-JUN-1989
; SEQ ID NO: 2
; LENGTH: 454
5260432-2

Query Match 70.8%; Score 34; DB 6; Length 454;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
||| |
DB 367 RRRPSQPY 374

RESULT 41
US-09-091-432-2
; Sequence 2, Application US/09091432
; Patent No. 5981837
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
; FILE REFERENCE: 7024-325
; CURRENT APPLICATION NUMBER: US/09/091,432
; CURRENT FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: PCT/US96/20094
; EARLIER FILING DATE: 1996-12-19
; EARLIER APPLICATION NUMBER: US 60/009,119
; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: US 60/013,388
; EARLIER FILING DATE: 1996-03-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Microsoft Word 2.0C
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sequence is deduced from DNA sequence of SEQ ID NO:1
; Patent No. 5981837
US-09-091-432-2

Query Match 70.8%; Score 34; DB 2; Length 520;

Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPRPY 8
||| |
DB 34 RRRPRPY 40

RESULT 42
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2

Query Match 70.8%; Score 34; DB 4; Length 2293;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 6
||| |
DB 2110 RRRPRP 2115

RESULT 43
US-09-264-419C-2
; Sequence 2, Application US/09264419C
; Patent No. 6174682
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: No. 6174682el Rhioedoxin Family Active Site Molecules and Us
; FILE REFERENCE: MNI-076
; CURRENT APPLICATION NUMBER: US/09/264,419C
; CURRENT FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-419C-2

Query Match 69.8%; Score 33.5; DB 4; Length 280;
Best Local Similarity 77.8%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

OY 1 RRRPRP-PY 8
||| |
DB 210 RRRPRPY 218

RESULT 44
US-09-024-975-9
; Sequence 9, Application US/09024975
; Patent No. 6133233
; GENERAL INFORMATION:

APPLICANT: ROSS, CHRISTOPHER R.
APPLICANT: BLECHA, FRANK
APPLICANT: SHI, JISHU
TITLE OF INVENTION: PEPTIDE MODULATION OF REPERFUSION INJURY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 2405 GRAND BLVD., SUITE 400
CITY: KANSAS CITY
STATE: MO
COUNTRY: USA
ZIP: 64108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/024,975
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/802,306
FILING DATE: 18-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: COLLINS, JOHN M.
REGISTRATION NUMBER: 26,262
REFERENCE/DOCKET NUMBER: 25585-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 816/474-9050
TELEFAX: 816/474-9057
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-024-975-9

Query Match 68.8%; Score 33; DB 4; Length 20;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRPP 7
1 |11111
Db 2 RRRRPP 8

RESULT 45
PCT-US95-12080-3
Sequence 3, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporaton
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea U. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-3

Query Match 68.8%; Score 33; DB 5; Length 59;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRPP 7
1 |11111
Db 2 RRRRPP 8

Search completed: September 24, 2001, 10:06:42
Job time: 82 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:06:05 : Search time 29.48 seconds
(without alignments)
20.672 Million cell updates/sec

Title: US-09-276-868-5

Perfect score: 48

Sequence: 1 RRRRRPPY 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	172	2 S68232	antimicrobial prot
2	41	85.4	104	2 S68607	hypothetical prote
3	41	85.4	372	2 S69501	DNA-binding prote
4	40	83.3	439	2 I60916	HNF-3/forkhead hom
5	39	81.2	488	2 T33739	hypothetical prote
6	38	79.2	82	2 A41051	spore coat protein
7	38	79.2	237	2 T43472	hypothetical prote
8	38	79.2	322	2 T35646	probable thiamin m
9	38	79.2	533	2 S75536	NADH dehydrogenase
10	38	79.2	700	2 F84601	probable ubiquitin
11	37	77.1	148	1 Q08E20	BLRF2 protein - hu
12	37	77.1	224	2 G82780	conserved hypothet
13	37	77.1	276	2 B83161	probable short-cha
14	37	77.1	498	2 C69587	L-arabinose isomer
15	36	75.0	70	2 A96650	protein F2401.6 [1
16	36	75.0	129	2 T03861	glycine-rich prote
17	36	75.0	171	2 T25696	hypothetical prote
18	36	75.0	190	1 S44958	lmkb protein - Str
19	36	75.0	248	1 J01682	infected cell prot
20	36	75.0	358	1 WMB838	infected cell prot
21	36	75.0	502	2 S69331	natriuretic peptid
22	36	75.0	543	1 JC4070	protein kinase (EC
23	36	75.0	549	2 JC5926	secreted klotho pr
24	36	75.0	903	2 T00074	hypothetical prote
25	36	75.0	992	1 GNMV93	structural polypor
26	36	75.0	1012	2 JC5925	membrane klotho pr
27	36	75.0	1063	1 GNMV94	structural polypor
28	36	75.0	1063	1 GNMV94	structural polypor
29	36	75.0	1063	1 GNMV77	structural polypor

30	36	75.0	2712	2 T30949	hypothetical prote
31	35	72.9	80	2 S03560	transition protein
32	35	72.9	130	2 S75341	hypothetical prote
33	35	72.9	130	2 C72532	hypothetical prote
34	35	72.9	200	2 S54834	Hsp peptide - huma
35	35	72.9	237	2 S28560	early light-induce
36	35	72.9	241	2 A75268	conserved hypothet
37	35	72.9	281	2 F83923	manganese-coplati
38	35	72.9	454	2 S06124	retinoic acid rece
39	35	72.9	611	2 C82881	hypothetical prote
40	35	72.9	638	2 S41159	sodium transport p
41	35	72.9	682	2 F83228	hypothetical prote
42	35	72.9	698	1 JX0202	long-chain-fatty-a
43	35	72.9	815	2 B82746	DNA topoisomerase
44	35	72.9	816	2 B86285	hypothetical prote
45	35	72.9	825	2 T27852	hypothetical prote

ALIGNMENTS

RESULT 1
S68232
antimicrobial protein PR-39 precursor, cathelin-associated - pig
N:Alliterate names: myeloid antibacterial protein PR-39
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68232; JN0899; I47138; S19563
R:Zhao, C.; Ganz, T.; Lehrer, R.I.
FEBS Lett. 376, 130-134, 1995
A>Title: Structures of genes for two cathelin-associated antimicrobial peptides: prop
A:Reference number: S68232; MUID:96105365
A:Accession: S68232
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-172 <ZHA>
A:Cross-references: EMBL:X89201; NID:g1165150; PIDN:CA61487.1; PID:g1165151
A:Title: A cDNA derived from pig bone marrow cells predicts a sequence identical to t
A:Reference number: JN0899; MUID:94071853
A:Accession: JN0899
A:Molecule type: mRNA
A:Residues: 1-20, 'A', 22-172 <STO>
A:Cross-references: GB:I23825; NID:9435100; PIDN:AAA31109.1; PID:9435101
A:Experimental source: bone marrow cells
R:Gudmundsson, G.H.; Magnusson, K.P.; Chowdhary, B.P.; Johansson, M.; Andersson, L.;
Proc. Natl. Acad. Sci. U.S.A. 92, 7085-7089, 1995
A>Title: Structure of the gene for porcine peptide antibiotic PR-39, a cathelin gene
A:Reference number: I47138; MUID:95350216
A:Accession: I47138
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28, 'T', 30-89, 'QR', 92-116, 'NDP', 120-172 <GUD>
A:Cross-references: EMBL:X87236; NID:9829142; PIDN:CA60682.1; PID:g1051298
R:Agelberth, B.; Lee, J.Y.; Bergman, T.; Carlquist, M.; Boman, H.G.; Mutt, V.; Joernv
Eur. J. Biochem. 202, 849-854, 1991
A:Title: Amino acid sequence of PR-39. Isolation from pig intestine of a new member o
A:Reference number: S19563; MUID:92111534
A:Accession: S19563
A:Molecule type: protein
A:Residues: 131-169 <AGR>
A:Experimental source: Intestine
C:Genetics:
A:Gene: PR39
A:introns: 66/3; 102/3; 126/3
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end; antibacterial
F:1-29/Domain: signal sequence #Status predicted <SIG>
F:122-129/Domain: cystatin homology <CTS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-169/Product: antimicrobial protein PR-39 #status experimental <MAT>

F:169/Modified site: amidated carboxyl end (Pro) (amide in mature form from following g)

Query Match 100.0%; Score 48; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPPY 8
131 RRRRPPY 138

RESULT 2

hypotheoretical protein At2g22000 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84607
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: G84607

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <STO>
A:Cross-references: GB:AE002093; NID:g4417292; PIDN:AA020417.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g22000
A:Map position: 2

Query Match 85.4%; Score 41; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPP 7
87 RRRRPP 93

RESULT 3

DNA-binding protein A variant - human

N:Alternate names: cold shock domain protein A
C:Species: Homo sapiens (man)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S69501
R:Coles, L.S.; Diamond, P.; Occhiodoro, F.; Vadas, M.A.; Shannon, M.F.
Nucleic Acids Res. 24, 2311-2317, 1996
A:Title: Cold shock domain proteins repress transcription from the GM-CSF promoter.
A:Reference number: S69501; MUID:96279731
A:Accession: S69501
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-372 <COI>
A:Cross-references: EMBL:X95325; NID:g1167837; PIDN:CA64631.1; PID:g1167838
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
C:Superfamily: Y box-binding protein 1; cold shock domain homology
F:93-157/Domain: cold shock domain homology <CSD>

Query Match 85.4%; Score 41; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPP 7
336 RRRRPP 342

RESULT 4

HNF-3/forkhead homolog-1 - rat

C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I60916
R:Clevidence, D.E.; Overdier, D.G.; Tao, W.; Qian, X.; Paul, L.; Lai, E.; Costa, R.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 3948-3952, 1993
A:Title: Identification of nine tissue-specific transcription factors of the hepatocyte
A:Reference number: A47450; MUID:93248207
A:Accession: I60916
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-439 <RES>
A:Cross-references: GB:LI3201; NID:g951067; PIDN:AAA74561.1; PID:g550513
C:Genetics:
A:Gene: HNF-1
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:100-192/Domain: fork head DNA-binding domain homology <FRD>

Query Match 83.3%; Score 40; DB 2; Length 439;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRRPPY 8
97 RRRRPPY 103

RESULT 5

hypotheoretical protein K08B4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33739
R:Rolling, T.; Murray, J.; Antonio, B.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid K08B4.
A:Reference number: Z21395
A:Accession: T33739
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-488 <ROH>
A:Cross-references: EMBL:AF100663; PIDN:AAC68982.1; GSPDB:GN00022; CESP:K08B4.5
C:Genetics:
A:Gene: CESP:K08B4.5
A:Map position: 4
A:Introns: 222/1; 318/2; 342/2; 399/3; 409/3; 440/2

Query Match 81.2%; Score 39; DB 2; Length 488;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRRPPY 8
345 RRRRPPY 352

RESULT 6

spore coat protein precursor - Bacillus subtilis

C:Species: Bacillus subtilis
C>Date: 03-Apr-1992 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S04835; A41051; F69606
R:Aranson, A.I.; Song, H.T.; Bourne, N.
Mol. Microbiol. 3, 437-444, 1989
A:Title: Gene structure and precursor processing of a novel Bacillus subtilis spore c
A:Reference number: S04835; MUID:89313296
A:Accession: S04835
A:Molecule type: DNA

A:Residues: 'MNVHTPNLSIRNMVKGIRKAREVELL', 2-82 <AR2>
 A:Cross-references: EMBL:X13740; NID:g39864; PIDN:CAA2004.1; PID:g39865
 A:Experimental source: strain JH642
 A>Note: part of this sequence, including the amino end of the mature protein, was confir
 R:Boune, N.; Fitzdames, P.C.; Atkinson, A.I.
 J. Bacteriol. 173, 6618-6625, 1991
 A:Title: Structural and germination defects of *Bacillus subtilis* spores with altered cor
 A:Reference number: A41051, M01D:92011439
 A:Accession: A41051
 A:Molecule type: protein
 A:Residues: 'XX', 3-11 <B0U>
 A:Experimental source: strain JH642
 A>Note: The material sequenced was the larger of two isolated precursor forms; the amino
 A>Note: both the location of the transcription start site and peptide sequencing of the
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
 C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Chn
 A.: Ehrlich, S.D.; Emmerson, P.T.; Ertlan, K.D.; Erington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Gallizzi, A.; Gall
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauei
 Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terstra, P.; Togonoi, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; M01D:98044033
 A:Accession: F69606
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 'MNVHTPNLSIRNMVKGIRKAREVELL', 2-82 <KUN>
 A:Cross-references: GB:Z59110; GB:AL009126; NID:g2633472; PIDN:CB13066.1; PID:el183229;
 A:Experimental source: strain 168
 C:Comment: This structural protein is expressed during stage V of sporulation.
 C:Genetics:
 A:Gene: colt
 A:Start codon: TTG
 C:Keywords: sporulation
 F:1-15/Domain: propeptide #status experimental <PRO>
 F:20-82/Product: spore coat protein #status experimental <MAT>
 RESULT 7
 T43472
 Hypothetical protein DKFZp434f1728.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
 C:Accession: T43472
 R:Blocker, H.; Boeher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z22513
 A:Accession: T43472
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-237 <AAA>
 A:Cross-references: EMBL:AL133615
 A:Experimental source: adult testis; clone DKFZp434f1728
 C:Genetics:
 A>Note: DKFZp434f1728.1

Query Match 79.2%; Score 38; DB 2; Length 237;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRPP 7
 Db 224 RRRPRPP 230

RESULT 8
 T35646
 probable thiamin monophosphate kinase - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Feb-2000
 C:Accession: T35646
 R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
 submitted to the EMBL Data Library, December 1998
 A:Reference number: Z21585
 A:Accession: T35646
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-322 <MOR>
 A:Cross-references: EMBL:AL034447; PIDN:CAA22405.1; GSPDB:GN00070; SCOEDB:SC7A1.06
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC7A1.06
 C:superfamily: hydrogenase expression/formation protein htype

Query Match 79.2%; Score 38; DB 2; Length 322;
 Best Local Similarity 85.7%; Pred. No. 56;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPRPP 8
 Db 192 RRRPRPP 198

RESULT 9
 S75536
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain F homolog - *Synechocystis* sp. (s
 N:Alternate names: protein sl11221
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 16-Jun-2000
 C:Accession: S75536
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys
 S*.
 A:Reference number: S74322; M01D:97061201
 A:Accession: S75536
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-533 <KAN>
 A:Cross-references: EMBL:DB90911; GB:AB001339; NID:g1655083; PIDN:BA18097.1; PID:g165
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Genetics:
 A:Gene: hoxF
 C:superfamily: *Synechocystis* NADH dehydrogenase (ubiquinone) chain F homolog; NADH de
 C:Keywords: 4Fe-4S; flavoprotein; FMN; iron-sulfur protein; metalloprotein; NAD; oxid
 F:1-104/Region: Desulfoylidrio NADP-reducing hydrogenase chain B similarity
 F:153-528/Domain: NADH dehydrogenase (ubiquinone) I chain F homology <NDOF>
 F:171-180/Region: NAD binding motif
 F:283-298/Region: FMN binding motif
 F:460,463,466,506/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 79.2%; Score 38; DB 1; Length 533;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RRRPPY 8
 DB 307 RRRPPY 312

RESULT 10
 F84601

Probable ubiquitin activating enzyme [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84601
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M. Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
 Ems, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: F84601
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-700 <STO>
 A:Cross-references: GB:AE002093; NID:94567278; PIDN:AA023691.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2G21470
 A:Map position: 2

Query Match 79.2%; Score 38; DB 2; Length 700;
 Best Local Similarity 87.5%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPY 8
 DB 257 RRRPPY 264

RESULT 11
 COAB20

BLF2 protein - human herpesvirus 4 (strain B95-8)
 C:Species: human herpesvirus 4, Epstein-Barr virus
 C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
 C:Accession: H43041; A03761; S33007
 R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983
 A:Title: Sequence analysis of the 17 166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
 A:Reference number: A93065; MUID:85035713
 A:Accession: H43041
 A:Molecule type: DNA
 A:Residues: 1-148 <BAN>
 A:Cross-references: EMBL:V01555; NID:959074; PIDN:CAA24853.1; PID:G1334867
 R:Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H.
 Nature 310, 207-211, 1984
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A:Reference number: A03794; MUID:84270667
 A:Contents: annotation; protein coding region
 C:Superfamily: human herpesvirus 4 BLF2 protein

Query Match 77.1%; Score 37; DB 1; Length 148;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPY 7
 DB 43 QRRPPY 49

RESULT 12
 G82780
 conserved hypothetical protein XF0649 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: G82780

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: G82780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-224 <STM>
 A:Cross-references: GB:AE003909; GB:AE003849; NID:99105513; PIDN:AAF83459.1; GSPDB:GN
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carr
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
 J.D.; Jungella, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tsunako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0649

Query Match 77.1%; Score 37; DB 2; Length 224;
 Best Local Similarity 71.4%; Pred. No. 57;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPPY 8
 DB 213 RRRPPY 219

RESULT 13

Probable short-chain dehydrogenase PA3883 [imported] - Pseudomonas aeruginosa (strain
 B83161
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: B83161
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: B83161
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-276 <STO>
 A:Cross-references: GB:AE004805; GB:AE004091; NID:99950055; PIDN:AMG07270.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3883
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 77.1%; Score 37; DB 2; Length 276;
 Best Local Similarity 85.7%; Pred. No. 69;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPPY 7
 DB 233 QRRPPY 239

RESULT 14
 C69587
 L-arabinose isomerase (EC 5.3.1.4) - Bacillus subtilis

C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: C69587
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertez, C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konlingstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Saito, T.; Scanlon, A.; Authors: Schlach, S.; Schreier, T.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serot, A.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. T. Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69587
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-498 <KUN>
 A:Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14840.1; PID:92635345
 A:Experimental source: strain 168
 C:Genetics: araA
 C:Function: araA
 A:Description: catalyzes the conversion of L-arabinose to L-ribulose, the first step in C:Superfamily: L-arabinose isomerase
 C:Keywords: arabinose metabolism; intramolecular oxidoreductase; isomerase

Query Match 77.18; Score 37; DB 2; Length 498;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

CY 1 RRRPPPY 8
 :||| |||
 Db 90 KRRPPPY 97

RESULT 15
 A96650
 protein F2401.6 [imported] - Arabidopsis thaliana
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96650
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L. Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Talion, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719
 A:Accession: A96650
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-70 <STO>
 A:Cross-references: GB:AE005173; NID:9740276; PIDN:AAF70835.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F2401.6
 A:Map position: 1

Query Match 75.08; Score 36; DB 2; Length 70;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 RRRPPPY 8
 :||| |||
 Db 64 RRRPPPY 70

RESULT 16
 T03861
 glycine-rich protein 15 - common tobacco (fragment)
 N:Alternate names: probable pollen wall protein
 C:Species: *Nicotiana tabacum* (common tobacco)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jan-2000
 C:Accession: T03861
 R:Fuenteberg, S.I.; Buccigaglia, P.A.; Smith, A.G.
 A:Description: Molecular characterization of an anther-specific gene from tobacco sho
 A:Reference number: Z15123
 A:Accession: T03861
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-129 <FUE>
 A:Cross-references: EMBL:U74325; NID:91658178; PIDN:AB18261.1; PID:91658179
 C:Genetics: GPR15
 C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 75.08; Score 36; DB 2; Length 129;
 Best Local Similarity 85.7%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 RRRPPPY 8
 :||| |||
 Db 28 RRRPPPY 34

RESULT 17
 T25696
 hypothetical protein F16F9.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T25696
 R:Fulton, B.
 A:Description: The sequence of C. elegans cosmid F16F9.
 A:Reference number: Z20071
 A:Accession: T25696
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-171 <FUL>
 A:Cross-references: EMBL:U67956; PIDN:AB07690.1; GSPDB:GN00028; CESP:F16F9.1
 A:Experimental source: strain Bristol N2; clone F16F9
 C:Genetics:
 A:Gene: CESP:F16F9.1
 A:Map position: X
 A:Introns: 28/3; 97/3; 111/3

Query Match 75.08; Score 36; DB 2; Length 171;
 Best Local Similarity 85.7%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 RRRPPPY 8
 :||| |||
 Db 12 RRRPPPY 18

RESULT 18
 S44958
 Imk protein - *Streptomyces lincolnensis*
 C:Species: *Streptomyces lincolnensis*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S69820; S44958
 R:Reschke, U.; Schmidt, H.; Zhang, H.Z.; Plepersberg, W.

Mol. Microbiol. 16, 1137-1156, 1995
A:Title: Molecular characterization of the lincomycin-production gene cluster of Streptococcus pneumoniae
A:Reference number: S69805; MUID:96020646
A:Accession: S69820
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-190 <PE2>
A:Cross-references: EMBL:X79146; NID:9499194; PID:9487701
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994
C:Genetics:
A:Gene: lmbK
C:Superfamily: lmbK protein

Query Match 75.0%; Score 36; DB 1; Length 190;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPY 8
||| |||
Db 164 RRRPPY 170

RESULT 19
JQ1682
Infected cell protein ICP34.5 - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 26-Feb-1999
C:Accession: JQ1682
R:Dolan, A.; McKie, E.; Maclean, A.R.; McGeoch, D.J.
J. Gen. Virol. 73, 971-973, 1992
A:Title: Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17.
A:Reference number: JQ1682; MUID:92341080
A:Accession: JQ1682
A:Molecule type: DNA
A:Residues: 1-248 <DOL>
A:Cross-references: GB:S40593
C:Genetics:
A:Gene: ICP34.5
C:Superfamily: herpesvirus infected cell protein ICP34.5
C:Keywords: tandem repeat
F:161-175/Region: 3-residue repeats (A-T-P)

Query Match 75.0%; Score 36; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
Db 11 RRRPP 16

RESULT 20
WMB38
Infected cell protein ICP34.5 - human herpesvirus 1 (strain F)
C:Species: human herpesvirus 1
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A27768
R:Chou, J.; Roizman, B.
J. Virol. 57, 629-637, 1986
A:Title: The terminal a sequence of the herpes simplex virus genome contains the promote
A:Reference number: A27768; MUID:86115412
A:Accession: A27768
A:Molecule type: DNA
A:Residues: 1-358 <CHO>
A:Cross-references: GB:M12240; NID:9330122; PID:AAA5794.1; PID:9330123
C:Comment: This protein accumulates late in infection predominantly in the cytoplasm of
C:Genetics:
A:Gene: ICP34.5
C:Superfamily: herpesvirus infected cell protein ICP34.5
C:Keywords: tandem repeat
F:175-204/Region: 3-residue repeats (A-T-P)

Query Match 75.0%; Score 36; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 13e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
Db 26 RRRPP 31

RESULT 21
S69331
natriuretic peptide receptor C NPR-C - Japanese eel
C:Species: Anguilla japonica (Japanese eel)
C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C:Accession: S69331
R:Takahima, A.; Katafuchi, T.; Shibasaki, M.; Kashiwagi, M.; Hagiwara, H.; Takel, Y.
Eur. J. Biochem. 227, 673-680, 1995
A:Title: Cloning, properties, site-directed mutagenesis analysis of the subunit struc
A:Reference number: S69331; MUID:95172050
A:Accession: S69331
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-502 <TAK>
A:Cross-references: EMBL:D31928; NID:9667988; PIDN:BA06711.1; PID:91532269
C:Superfamily: natriuretic peptide receptor C; natriuretic peptide-binding domain hom
C:Keywords: receptor
F:96-428/Domain: natriuretic peptide-binding domain homology <NPB>

Query Match 75.0%; Score 36; DB 2; Length 502;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
|||||
Db 422 RRRPP 428

RESULT 22
JC4070
Protein kinase (EC 2.7.1.37) A - Streptomyces coelicolor (strain A3-2)
N:Alternate names: serine/threonine protein kinase A
C:Species: Streptomyces coelicolor
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JC4070
R:Urahe, H.; Ogawara, H.
Gene 153, 99-104, 1995
A:Title: Cloning, sequencing and expression of serine/threonine kinase-encoding genes
A:Reference number: JC4070; MUID:95189113
A:Accession: JC4070
A:Molecule type: DNA
A:Residues: 1-543 <URA>
A:Cross-references: GB:D66821; DDBJ:D26539; NID:91483185; PIDN:BA13168.1; PID:966602
C:Genetics:
A:Gene: pkaa
C:Superfamily: Streptomyces coelicolor protein kinase A; protein kinase homology
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase
F:6-275/Domain: protein kinase homology <KIN>

Query Match 75.0%; Score 36; DB 1; Length 543;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 7
|||||
Db 464 RRRPP 470

RESULT 23
JC5926

secreted klotho protein - human
C:Species: Homo sapiens (man)
C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
C:Accession: JCS926
R:Matsumura, Y.; Alizawa, H.; Shiraki-Tida, T.; Nagai, R.; Kuro-o, M.; Nabeshima, Y.
Biochem. Biophys. Res. Commun. 242, 626-630, 1998
A:Title: Identification of the human klotho gene and its two transcripts encoding membra
A:Reference number: JCS925; MUID:98125543
A:Accession: JCS926
A:Molecule type: DNA
A:Residues: 1-349 <MAT>
A:Cross-references: DDBJ:AB009666
A:Note: neither the complete nucleic acid sequence nor the complete translation are show
C:Comment: This protein plays a part in a signaling pathway that regulates aging and the
C:Genetics:
A:Gene: kl
A:Map position: 13q12
C:Superfamily: Agrobacterium beta-glucosidase

Query Match 75.0%; Score 36; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRP RP 7
|||||
DB 8 RRP RP 13

RESULT 24
T00074
Hypothetical protein KIAA0460 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00074
R:Seki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;
DNA Res. 4, 345-349, 1997
A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A:Reference number: Z14085; MUID:98116662
A:Accession: T00074
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-903 <SEK>
A:Cross-references: EMBL:AB007929; NID:g3413881; PIDN:BAA32305.1; PID:g3413882
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0460

Query Match 75.0%; Score 36; DB 2; Length 903;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRP RP 8
|||||
DB 879 RRP RP 885

RESULT 25
GMMVRA3
structural polyprotein - rubella virus (strain M33)
N:Contains: capsid protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: rubella virus
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A27505
R:Clarke, D.M.; Loo, T.W.; Hui, I.; Chong, P.; Gillam, S.
Nucleic Acids Res. 15, 3041-3057, 1987
A:Title: Nucleotide sequence and in vitro expression of rubella virus 24S subgenomic mes
A:Reference number: A27505; MUID:87174825
A:Accession: A27505
A:Molecule type: mRNA
A:Residues: 1-992 <GLA>
A:Cross-references: GB:X05259; NID:g62093; PIDN:CAA28880.1; PID:g62094

C:Superfamily: rubella virus structural polyprotein
C:Keywords: capsid protein; glycoprotein; polyprotein; transmembrane protein
F:1-299/Product: capsid protein #status predicted <NCP>
F:277-297/Domain: transmembrane #status predicted <TM1>
F:300-580/Product: membrane glycoprotein E2 #status predicted <ME2>
F:410-428/Domain: transmembrane #status predicted <TM2>
F:515-533/Domain: transmembrane #status predicted <TM3>
F:535-553/Domain: transmembrane #status predicted <TM4>
F:581-992/Product: membrane glycoprotein E1 #status predicted <ME1>
F:583-601/Domain: transmembrane #status predicted <TM5>
F:664-682/Domain: transmembrane #status predicted <TM6>
F:965-983/Domain: transmembrane #status predicted <TM7>
F:352,370,428,656,757,789/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 75.0%; Score 36; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRP RP 7
|||||
DB 35 RRP RP 40

RESULT 26
JCS925
membrane klotho protein - human
C:Species: Homo sapiens (man)
C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-1999
C:Accession: JCS925
R:Matsumura, Y.; Alizawa, H.; Shiraki-Tida, T.; Nagai, R.; Kuro-o, M.; Nabeshima, Y.
Biochem. Biophys. Res. Commun. 242, 626-630, 1998
A:Title: Identification of the human klotho gene and its two transcripts encoding mem
A:Reference number: JCS925; MUID:98125543
A:Accession: JCS925
A:Molecule type: DNA
A:Residues: 1-1012 <MAT>
A:Cross-references: DDBJ:AB009666
A:Note: neither the complete nucleic acid sequence nor the complete translation are s
C:Comment: This protein plays a part in a signaling pathway that regulates aging and
C:Genetics:
A:Gene: kl
A:Map position: 13q12

Query Match 75.0%; Score 36; DB 2; Length 1012;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRP RP 7
|||||
DB 8 RRP RP 13

RESULT 27
GMMVRA4
structural polyprotein - rubella virus (strain Thierien)
N:Contains: capsid protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: rubella virus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jul-2000
C:Accession: A29822; A26884; B29811; B35320; A32270; B32270
R:Takkinen, K.; Vildgren, G.; Ekstrand, J.; Hellman, U.; Kalkkinnen, N.; Wernstedt, C.;
J. Gen. Virol. 69, 603-612, 1988
A:Title: Nucleotide sequence of the rubella virus capsid protein gene reveals an unus
A:Reference number: A29822; MUID:88171449
A:Accession: A29822
A:Molecule type: genomic RNA
A:Residues: 1-319 <TAK>
A:Cross-references: GB:D00242; NID:g222578; PIDN:BAA00172.1; PID:g222579
R:Vildgren, G.; Takkinen, K.; Kalkkinnen, N.; Kaariainen, L.; Pettersson, R.F.
J. Gen. Virol. 68, 2347-2357, 1987
A:Title: Nucleotide sequence of the genes coding for the membrane glycoproteins E1 an
A:Reference number: A32270; MUID:88009948

A:Accession: A26884
A:Molecule type: genomic RNA
A:Residues: 268-1063 <VID>
A:Cross-references: GB:D00156; MID:9222576; PIDN:BA28178.1; PID:g3133179
R:Frey, T.K.; Marr, L.D.
Gene 62, 85-99, 1988
A:Title: Sequence of the region coding for virion proteins C and E2 and the carboxy term
A:Reference number: A29811; MUID:88226020
A:Accession: B29811
A:Molecule type: genomic RNA
A:Residues: 1-86, 'T', 88-162, 'E', 164-318, 'C', 320-394, 'D', 396-544, 'L', 546-621 <FRE>
A:Cross-references: GB:M18901
R:Dominguez, G.; Wang, C.Y.; Frey, T.K.
Virology 177, 225-238, 1990
A:Title: Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement
A:Reference number: A35320; MUID:90281585
A:Accession: B35320
A:Molecule type: genomic RNA
A:Residues: 1-86, 'T', 88-162, 'E', 164-318, 'C', 320-394, 'D', 396-544, 'L', 546-958, 'L', 960-990,
A:Cross-references: GB:M15240; GB:M32735; MID:9333971; PIDN:AAA8529.1; PID:g333973
C:Superfamily: rubella virus structural polyprotein
C:Keywords: capsid protein; glycoprotein; polyprotein; transmembrane protein
F:2-277/Product: capsid protein #status predicted <CAP>
F:278-562/Product: membrane glycoprotein E2 precursor #status predicted <PG2>
F:278-299/Domain: nonterminal signal sequence #status predicted <E2S>
F:300-562/Product: membrane glycoprotein E2 #status predicted <GL2>
F:539-555/Domain: transmembrane #status predicted <TM2>
F:563-1063/Product: membrane glycoprotein E1 precursor #status predicted <PG1>
F:563-1063/Product: nonterminal signal sequence #status predicted <E1S>
F:583-1063/Product: membrane glycoprotein E1 #status predicted <GL1>
F:1032-1050/Domain: transmembrane #status predicted <TM3>
F:353,371,410,429,658,759,791/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
Db 35 RRRPP 40

RESULT 28
GNMVA
structural polyprotein - rubella virus (strain RA27/3 vaccine)
N:Contains: capsid protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: rubella virus
C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 16-Jul-1999
C:Accession: S04800
R:Nakhasi, H.L.; Thomas, D.; Zheng, D.; Liu, T.Y.
Nucleic Acids Res. 17, 4393-4394, 1989
A:Title: Nucleotide sequence of capsid, E2 and E1 protein genes of rubella virus vaccine
A:Reference number: S04800; MUID:89296505
A:Accession: S04800
A:Molecule type: mRNA
A:Residues: 1-1063 <NAK>
A:Cross-references: GB:X14871; NID:61915; PIDN:CAA33016.1; PID:961916
A:Note: the authors translated the codon CCG for residue 207 as Val and AGC for residue
C:Superfamily: rubella virus structural polyprotein
C:Keywords: capsid protein; glycoprotein; polyprotein; transmembrane protein
F:1-300/Product: capsid protein C #status predicted <NPC>
F:278-299/Domain: transmembrane #status predicted <TM1>
F:301-583/Product: membrane glycoprotein E2 #status predicted <E2G>
F:539-555/Domain: transmembrane #status predicted <TM2>
F:583-1063/Product: membrane glycoprotein E1 #status predicted <E1G>
F:1032-1050/Domain: transmembrane #status predicted <TM3>
F:353,371,410,429,658,759,791/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
Db 35 RRRPP 40

RESULT 29
GNMVT7
structural polyprotein - rubella virus (strain HPV77)
N:Contains: capsid protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: rubella virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: J00087
R:Zheng, D.; Dickens, L.; Liu, T.Y.; Nakhasi, H.L.
Gene 82, 343-349, 1989
A:Title: Nucleotide sequence of the 24S subgenomic messenger RNA of a vaccine strain
A:Reference number: J00087; MUID:90060825
A:Accession: J00087
A:Molecule type: mRNA
A:Residues: 1-1063 <ZHE>
A:Cross-references: GB:M30776; MID:9333974; PIDN:AAA47421.1; PID:g333975
C:Superfamily: rubella virus structural polyprotein
C:Keywords: capsid protein; glycoprotein; polyprotein; transmembrane protein
F:1-300/Product: capsid protein #status predicted <CAP>
F:278-299/Domain: transmembrane #status predicted <TM1>
F:301-582/Product: membrane glycoprotein E2 #status predicted <E2S>
F:539-555/Domain: transmembrane #status predicted <TM2>
F:583-1063/Product: membrane glycoprotein E1 #status predicted <E1S>
F:1032-1050/Domain: transmembrane #status predicted <TM3>
F:353,371,429,658,759,791/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 75.0%; Score 36; DB 1; Length 1063;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
|||||
Db 35 RRRPP 40

RESULT 30
T30949
hypothetical protein C44E4.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30949
R:Sammons, L.; Wohldmann, P.; Gilliam, B.
Submitted to the EMBL Data Library, August 1999
A:Description: The sequence of C. elegans cosmid C44E4.
A:Reference number: Z20945
A:Accession: T30949
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2712 <SAM>
A:Cross-references: EMBL:AF003140; PIDN:AD47119.1
A:Experimental source: strain Bristol N2; clone C44E4
C:Genetics:
A:Map position: 1
A:Intons: 44/1; 346/2; 870/1; 1005/3; 1404/3; 1506/2; 1665/1; 1891/3; 1936/3; 2068/2
A:Note: C44E4.1a

Query Match 75.0%; Score 36; DB 2; Length 2712;
Best Local Similarity 75.0%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPP 8
|||||
Db 1413 RRRPP 1420

RESULT 31

503560
transition protein S2, spermatid - smaller spotted catshark
N:Alternate names: basic nuclear protein S2
C:Species: Scyllorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Feb-1997
C:Accession: S03560
R:Chavliere, M.; Martiague, A.; Briand, G.; Sautiere, P.; Chevallier, P.
Eur. J. Biochem. 180, 329-335, 1989
A:Title: Nuclear basic protein transition during sperm differentiation. Primary structure
A:Reference number: S03560; MUID:89170733
A:Accession: S03560
A:Molecule type: protein
A:Residues: 1-80 <CNA>
C:Keywords: DNA binding

Query Match 72.9%; Score 35; DB 2; Length 80;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
 |||||
Db 26 KRRPKPP 32

RESULT 32
S75341
hypothetical protein slr2010 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S75341
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
K. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201
A:Accession: S75341
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <KAN>
A:Cross-references: EMBL:D90904; GB:AB001339; NID:G1652225; PIDN:BA17255.1; PID:dl01798
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 72.9%; Score 35; DB 2; Length 130;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPRPP 8
 |||||
Db 38 RRRPRPP 44

RESULT 33
C72532
hypothetical protein APE2231 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C72532
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339
A:Accession: C72532
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1190 <KAW>
A:Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BA81243.1; PID:dl045029; PID:9510
A:Experimental source: strain K1
C:Genetics:

A:Gene: APE2231

Query Match 72.9%; Score 35; DB 2; Length 190;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
 |||||
Db 61 RRRPRPP 67

RESULT 34
S54834
HP8 peptide - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 03-Nov-2000
C:Accession: S54834
R:Xu, L.; Man, D.; Liu, Y.; Li, H.; Zhang, P.; Sul, Y.; Gu, J.
submitted to the EMBL Data Library, May 1995
A:Description: HP8: the partial cDNA clone of a novel member of C/EBP gene family.
A:Reference number: S54834
A:Accession: S54834
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-200 <XUL>
A:Cross-references: EMBL:X87248
C:Superfamily: CCAT/enhancer-binding protein alpha

Query Match 72.9%; Score 35; DB 2; Length 200;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
 |||||
Db 101 RRRPRPP 107

RESULT 35
S28560
early light-induced protein - sensitive fern (fragment)
C:Species: Onoclea sensibilis (sensitive fern)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C:Accession: S28560
R:Kamalay, J. C.
submitted to the EMBL Data Library, November 1992
A:Reference number: S28560
A:Accession: S28560
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-237 <KAM>
A:Cross-references: EMBL:Z18809; NID:920197; PID:920198

Query Match 72.9%; Score 35; DB 2; Length 237;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPRPP 7
 |||||
Db 75 RRRPRPP 81

RESULT 36
A75268
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: A75268
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
M.; Shen, M.; Yamahewyan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: A75268
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <NR1>
A:Cross-references: GB:AE002078; GB:AE000513; NID:g6460306; PIDN:AAF12030.1; PID:g6460306
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2490
A:Map position: 1

Query Match 72.9%; Score 35; DB 2; Length 241;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPY 7
||| |
Db 77 RRRPAP 83

RESULT 37
F83923
manganese-containing catalase BH2190 [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: F83923
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20263314
A:Accession: F83923
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-781 <ST0>
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05909.1; GSPDB:GN04
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2190

Query Match 72.9%; Score 35; DB 2; Length 281;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPY 8
||| |
Db 258 RLRPAPY 265

RESULT 38
S06124
retinoic acid receptor delta - eastern newt
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Aug-1999
C:Accession: S06124; S09180
R:Ragsdale Jr., C.W.; Petkovich, M.; Gates, P.B.; Chambon, P.; Brookes, J.P.
Nature 341, 654-657, 1989
A:Title: Identification of a novel retinoic acid receptor in regenerative tissues of the
A:Reference number: S06123; MUID:90015190
A:Accession: S06124
A:Molecule type: mRNA
A:Residues: 1-454 <RAG>
A:Cross-references: EMBL:X17586; NID:g64134; PIDN:CMA35603.1; PID:g64135
R:Giguere, V.; Ong, E.S.; Evans, R.M.; Tablin, C.J.
Nature 337, 566-569, 1989
A:Title: Spatial and temporal expression of the retinoic acid receptor in the regenerati
A:Reference number: S02758; MUID:89127522
A:Accession: S09180
A:Molecule type: mRNA
A:Residues: 215-229, 'D', 231-251, 'S', 253-445, 'A', 447-454 <GIG>

C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F:90-341/Domain: erba transforming protein homology <ERBA>

Query Match 72.9%; Score 35; DB 2; Length 454;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPPY 8
||| |
Db 369 RRRPKPY 376

RESULT 39
C82881
hypothetical protein U0520 [Imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: C82881
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: C82881
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-611 <GLA>
A:Cross-references: GB:AE002150; GB:AF222894; NID:g6899515; PIDN:AAF30933.1; GSPDB:GN
C:Genetics:
A:Gene: U0520
A:Genetic code: SGC3
C:Superfamily: Ureaplasma urealyticum hypothetical protein U0520

Query Match 72.9%; Score 35; DB 2; Length 611;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPY 8
||| |
Db 519 RRORPPY 525

RESULT 40
S41159
sodium transport protein beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jun-2000
C:Accession: S41159
R:Canessa, C.M.; Schild, L.; Buell, G.; Thorens, B.; Gautschi, I.; Horiwanger, J.D.;
Nature 367, 463-467, 1994
A:Title: Amiloride-sensitive epithelial Na(+) channel is made of three homologous sub
A:Reference number: S41158; MUID:94150624
A:Accession: S41159
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-638 <CAN>
A:Cross-references: EMBL:X77932
C:Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I r

Query Match 72.9%; Score 35; DB 2; Length 638;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPY 8
||| |
Db 562 RRRORPY 569

RESULT 41

F83228
Hypothetical protein PA3340 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83228
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Micooguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardidg, K.; Lam,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950, MIMD:20437337
A:Accession: F83228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-682 <STO>
A:Cross-references: GB:AE004756; GB:AE004091; NID:9949466; PIDN:AG06728.1; GSPDB:GN001
C:Genetics:
A:Experimental source: strain PA01
A:Gene: PA3340

Query Match 72.9%; Score 35; DB 2; Length 682;
Best local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPP 7
|||
Db 396 RRRPRPP 402

RESULT 42
JX0202
Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JX0202
R:Abbe, T.; Fujino, T.; Fukuyama, R.; Minoshima, S.; Shimizu, N.; Toh, H.; Suzuki, H.; Ya
J. Biochem. 111, 123-128, 1992
A:Title: Human long-chain acyl-CoA synthetase: structure and chromosomal location.
A:Reference number: JX0202; MIMD:92299644
A:Accession: JX0202
A:Molecule type: mRNA
A:Residues: 1-698 <ABE>
A:Cross-references: GB:D10040; NID:9219899; PIDN:BA00931.1; PID:9219900
A:Experimental source: liver
C:Genetics:
A:Gene: GDB:FAC12
A:Cross-references: GDB:127357; OMIM:152426
A:Map position: 4q34-q35
C:Superfamily: human long-chain-fatty-acid-CoA ligase; acetate-CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; hydrolase
F:143-681/Domain: acetate-CoA ligase homology <ACD>

Query Match 72.9%; Score 35; DB 1; Length 698;
Best local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRPP 7
|||
Db 259 RRRPRPP 265

RESULT 43
B82746
DNA topoisomerase I XF0920 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: B82746
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MIMD:20365717

A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82746
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-815 <SIM>
A:Cross-references: GB:AE003931; GB:AE003849; NID:99105834; PIDN:AAF83730.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Fachinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferio, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeti,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0920
C:Superfamily: bacterial type I DNA topoisomerase

Query Match 72.9%; Score 35; DB 2; Length 815;
Best local Similarity 62.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
:||||
Db 257 KRRPRPPF 264

RESULT 44
B86285
Hypothetical protein AAD39642.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86285
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MIMD:21016719
A:Accession: B86285
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-816 <STO>
A:Cross-references: GB:AE005172; NID:95103812; PIDN:AA039642.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 72.9%; Score 35; DB 2; Length 816;
Best local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPRPPY 8
:||||
Db 792 QRRPRPPY 798

RESULT 45
T27852
Hypothetical protein ZK418.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T27852
 R;Fulton, L.
 submitted to the EMBL Data Library, April 1994
 A;Description: The sequence of C. elegans cosmid ZK418.
 A;Reference number: 220430
 A;Accession: T27852
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-825 <FULL>
 A;Cross-references: EMBL:U00047; PIDN:AA50690.1; CESP:ZK418.6
 A;Experimental source: strain Bristol N2
 C;Genetics:
 A;Gene: CESP:ZK418.6
 A;Introns: 24/3; 53/3; 98/2; 153/1; 196/3; 269/1; 338/2; 394/2; 444/1; 562/2; 644/3

Query Match 72.9%; Score 35; DB 2; Length 825;
 Best Local Similarity 85.7%; Pred No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRRP 7
 |||||
 Db 464 RRRPAP 470

Search completed: September 24, 2001, 10:06:06
 Job time: 46 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2001, 10:09:06 : Search time 16.18 Seconds
(without alignments)
16.937 Million cell updates/sec

Title: US-09-276-868-5
Perfect score: 48
Sequence: 1 RRRPRPV 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_39.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	172	1 PR39_PIG	P80054 sus scrofa
2	41	85.4	372	1 DBPA_HUMAN	P16089 homo sapien
3	41	85.4	625	1 DUS8_HUMAN	O13202 homo sapien
4	40	83.3	439	1 HPH1_RAT	O63244 rattus norv
5	38	79.2	107	1 COTT_BACSU	P11863 bacillus su
6	37	77.1	148	1 YLL2_EBV	P03199 Epstein-Bar
7	37	77.1	148	1 YLL2_EBV	O07285 Epstein-Bar
8	37	77.1	498	1 ARAA_BACSU	P94523 bacillus su
9	36	75.0	245	1 ICP3_HSV1N	P37319 herpes simp
10	36	75.0	248	1 ICP3_HSV11	P36313 herpes simp
11	36	75.0	252	1 ICP3_HSV1D	P37318 herpes simp
12	36	75.0	263	1 ICP3_HSV1F	P08353 herpes simp
13	36	75.0	333	1 CIW4_HUMAN	O9HY98 homo sapien
14	36	75.0	543	1 PKAA_STRCO	P54739 streptomyce
15	36	75.0	638	1 SCAB_MOUSE	O9WU38 mus musculu
16	36	75.0	638	1 SCAB_RAT	P37090 rattus norv
17	36	75.0	992	1 POLS_RUBVM	P08563 rubella vir
18	36	75.0	1063	1 POLS_RUBVR	P21480 rubella vir
19	36	75.0	1063	1 POLS_RUBVR	P19725 rubella vir
20	36	75.0	1063	1 POLS_RUBVT	P07566 rubella vir
21	35	72.9	80	1 SSS2_SCYCA	P11020 scyllorhinu
22	35	72.9	505	1 RRG_MOTVI	P18516 notophtham
23	35	72.9	698	1 LCFB_HUMAN	P33121 homo sapien
24	35	72.9	699	1 LCFB_HUMAN	P41215 homo sapien
25	35	72.9	2109	1 PKS1_ASPPA	O12053 aspergillus
26	34	70.8	15	1 MK1_PALPR	P80408 palomona pr
27	34	70.8	180	1 XG_HUMAN	P55808 homo sapien
28	34	70.8	215	1 Y215_ADE02	P02291 human adeno
29	34	70.8	261	1 RL1_HSV2H	P28283 herpes simp
30	34	70.8	264	1 RS31_ARATV	P05916 arbidopsin
31	34	70.8	333	1 HME1_CHICK	O67764 rattus norv
32	34	70.8	361	1 DBPA_RAT	P22932 homo sapien
33	34	70.8	443	1 RRG2_HUMAN	

ALIGNMENTS

RESULT ID	PR39_PIG	STANDARD	PRT	172 AA.
AC	P80054; O9TR84;			
DT	01-MAR-1992 (Rel. 21, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ANTIBACTERIAL PROTEIN PR-39 PRECURSOR.			
GN	PR39.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95350216; PubMed=7624374;			
RA	Gudmundsson G.H., Magnusson R.P., Chowdhary B.P., Johansson M.,			
RA	Andersson L., Boman H.G.;			
RT	"Structure of the gene for porcine peptide antibiotic PR-39, a			
RT	cathelin gene family member: comparative mapping of the locus for the			
RT	human peptide antibiotic FAL-39."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:7085-7089(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=94071853; PubMed=8250863;			
RA	Storici P., Zanetti M.;			
RT	"A cDNA derived from pig bone marrow cells predicts a sequence			
RT	identical to the intestinal antibacterial peptide PR-39."			
RL	Biochem. Biophys. Res. Commun. 196:1058-1065(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=96105365; PubMed=7498526;			
RA	Zhao C., Ganz T., Lehrer R.I.;			
RT	"Structures of genes for two cathelin-associated antimicrobial			
RT	peptides: prophelin-2 and PR-39."			
RL	FEBS Lett. 376:130-134(1995).			
RN	[4]			
RP	SEQUENCE OF 131-169.			
RC	TISSUE=Intestine;			
RX	MEDLINE=92111534; PubMed=1765098;			
RA	Aggerholm B., Lee J.-Y., Bergman T., Carlquist M., Boman H.G.,			
RT	Mott V., Joernvall H.;			
RT	"Amino acid sequence of PR-39. Isolation from pig intestine of a new			
RT	member of the family of proline-arginine-rich antibacterial			
RL	peptides."			
RL	Eur. J. Biochem. 202:849-854(1991).			
RN	[5]			
RP	SEQUENCE OF 131-164, AND FUNCTION.			
RC	TISSUE=Neutrophils;			
RX	MEDLINE=95088504; PubMed=7996056;			
RA	Shi J., Ross C.R., Chengappa M.M., Blecha F.;			
RT	"Identification of a proline-arginine-rich antibacterial peptide from			
RT	neutrophils that is analogous to PR-39, an antibacterial peptide from			
RT	the small intestine."			

34	34	70.8	454	1	RRG1_HUMAN	P13631 homo sapien
35	34	70.8	520	1	CP84_ARATV	O42600 arbidopsin
36	34	70.8	520	1	GAG_SIVAI	P27972 simian immu
37	34	70.8	521	1	GAG_SIVAG	P27978 simian immu
38	34	70.8	587	1	YNJ4_YEAST	P53993 saccharomyc
39	34	70.8	639	1	YGFT_ECOLI	O46820 escherichia
40	34	70.8	801	1	BRD2_HUMAN	P25440 homo sapien
41	34	70.8	834	1	5HTA_DROME	P28285 drosophila
42	34	70.8	936	1	ORP1_MOUSE	P56716 mus musculu
43	34	70.8	2052	1	YDOB_SCHPO	O13731 schizosacch
44	34	70.8	2156	1	ORP1_HUMAN	P36715 homo sapien
45	33	68.8	101	1	E311_ADE02	P24935 human adeno

CN DUSP8 OR VHS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96009533; PubMed-7561881;
 RT "Marelli K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.,
 RT "Hv-5: a protein tyrosine phosphatase abundant in brain that
 RT inactivates mitogen-activated protein kinase.";
 RL J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
 CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: U27193; AAA83151.1; -
 DR MIM: 602038; -
 DR InterPro: IPR000340; -
 DR InterPro: IPR000387; -
 DR InterPro: IPR002965; -
 DR Pfam: PF00782; DSCP; 1.
 DR PRINTS: PR01217; PRICEXTENS.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 KW Hydrolyase; Nuclear protein.
 FT DOMAIN 28 43 CH2 A DOMAIN.
 FT DOMAIN 117 132 CH2 B DOMAIN.
 FT DOMAIN 162 430 CATALYTIC.
 FT DOMAIN 310 550 PRO-RICH.
 FT ACT_SITE 246 246 BY SIMILARITY.
 DE SEQUENCE 625 AA; 65840 MW; DCBEA14487219666 CRC64;
 SJ
 Query Match 85.4%; Score 41; DB 1; Length 625;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRRPP 7
 DB 454 RRRRPP 460
 ID HFL_RAT
 AC 063244; STANDARD; PRT; 439 AA.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HEPYOCYTE NUCLEAR FACTOR 3 FORKHEAD HOMOLOG 1 (HFH-1).
 GN HFH1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID-10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Lung;
 RX MEDLINE-93248207; PubMed-7683413;
 RA Cleveland D.E., Overdier D.G., Tao W., Qian X., Paul L., Lai E.,
 RA Costa R.H.;
 RT "Identification of nine tissue-specific transcription factors of the
 RT hepatocyte nuclear factor 3/forkhead DNA-binding-domain family";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3948-3952(1993).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: L13201; AAA74561.1; -
 DR TRANSFAC: T02288; -
 DR InterPro: IPR001766; -
 DR Pfam: PF00250; Fork_head; 1.
 DR PRINTS: PR00053; FORKHEAD.
 DR PROSITE: PS00657; FORKHEAD_1; 1.
 DR PROSITE: PS00658; FORKHEAD_2; 1.
 DR PROSITE: PS50039; FORKHEAD_3; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation.
 FT DNA_BIND 99 190 FORK-HEAD.
 DE SEQUENCE 439 AA; 45481 MW; A37028448644C350 CRC64;
 SJ
 Query Match 83.3%; Score 40; DB 1; Length 439;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RRRPPY 8
 DB 97 RRRPPY 103
 ID COTT_BACSU
 AC P11663; STANDARD; PRT; 107 AA.
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE SPORE COAT PROTEIN T PRECURSOR.
 GN COTT.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 NC NCB1_TaxID-1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-168 / JH642;
 RX MEDLINE-89313286; PubMed-2546006;
 RA Atkinson A.I., Song H.Y., Bourne N.;
 RT "Gene structure and precursor processing of a novel Bacillus subtilis
 RT spore coat protein.";
 RL Mol. Microbiol. 3:437-444(1989).
 CC -1- FUNCTION: POSSIBLY PROTECTION OF SPORE AND PROBABLY PLAYS
 CC SOME ROLE IN GERMINATION.
 CC -1- SUBCELLULAR LOCATION: OUTER SURFACE OF ENDOSPORE.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

DR EMBL: X13740: CAA32004.1; -
DR EMBL: Z99110: CAB13066.1; -
DR PIR: S04835; S04835.
DR Subtilisin; Bg10495; coct.
KW Sporulation; Signal.
FT SIGNAL 1 44
FT CHAIN 45 107
SQ SEQUENCE 107 AA; 12992 MW; AD1F66F0C4CE29A3 CRC64;

Query Match
Best Local Similarity 79.2%; Score 38; DB 1; Length 107;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRRPP 8
    |||||
DB 82 RRRPP 87

RESULT 6
YLL2_EBV STANDARD; PRT; 148 AA.
ID YLL2_EBV P03199;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE HYPOTHETICAL BLF2 PROTEIN.
GN BLF2;
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
NCBI_TaxID=10377;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=64270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrell B.G.;
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RA Nature 310:207-211(1984).
RL Mature 310:207-211(1984).
CC -! CAUTION: BLF2 IS KNOWN AS BLF3 IN REF.1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: V01555; CAA24853.1; -
DR PIR: A03761; Q0BE20.
DR PIR: S33007; S33007.
KW Hypothetical protein; Early protein.
SQ SEQUENCE 148 AA; 16651 MW; AC4B184F0BCB391 CRC64;

Query Match
Best Local Similarity 77.1%; Score 37; DB 1; Length 148;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPP 7
    |||||
DB 43 QRRPP 49

RESULT 7
YLL2_EBV8 STANDARD; PRT; 148 AA.
ID YLL2_EBV8
AC 007285;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

```

```

DE HYPOTHETICAL BLF2 PROTEIN.
GN BLF2.
OS Epstein-Barr virus (strain AG876) (Human herpesvirus 4), and
OS Epstein-Barr virus (strain P3HR-1) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
NCBI_TaxID=82830, 82829;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=AG876, AND P3HR-1;
RX MEDLINE=93331716; PubMed=8393237;
RA Lees J.F., Arrand J.E., Pepper S.V., Stewart J.P., Mackett M.,
RA Arrand J.R.;
RT "The Epstein-Barr virus candidate vaccine antigen gp340/220 is highly
RT conserved between virus types A and B.";
RL Virology 195:578-586(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: L07922; AAA02784.1; -
DR EMBL: L07923; AAA02788.1; -
KW Hypothetical protein; Early protein.
SQ SEQUENCE 148 AA; 16669 MW; AC4B0953BC8A54A CRC64;

Query Match
Best Local Similarity 77.1%; Score 37; DB 1; Length 148;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPP 7
    |||||
DB 43 QRRPP 49

RESULT 8
ARAA_BACSU STANDARD; PRT; 498 AA.
ID ARAA_BACSU
AC P94523;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE L-ARABINOSIDE ISOMERASE (EC 5.3.1.4).
GN ARAA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=97124191; PubMed=8969504;
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,
RA Sanders J., Emerson P.T., Harwood C.R.;
RA "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress
RT responses, the utilization of plant cell walls and primary
RT metabolism.";
RL Microbiology 142:3067-3078(1996).
CC -! CATALYTIC ACTIVITY: L-ARABINOSIDE - L-RIBULOSE.
CC -! PATHWAY: FIRST STEP OF L-ARABINOSIDE CATABOLISM AFTER IT IS
CC TRANSPORTED INTO THE CELL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; 275208; CAA9587.1; -

CC DR EMBL; 299118; CAB14840.1; -

CC DR Subtilist; BG11904; araA.

CC KM Arabidose catabolism; Isomerase.

SO SEQUENCE 498 AA; 56540 MW; D6D50EB90514494C CRC64;

Query Match 77.1%; Score 37; DB 1; Length 498;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 RRRPPPP 8
:111111111
Db 90 KRRPPPP 97

RESULT 9
ICP3_HSV1N STANDARD; PRT; 245 AA.

ID ICP3_HSV1N P37319;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
GN ICP34.5.
CS Herpes simplex virus (type 1 / strain MGH-10).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=37107;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90156494; PubMed=2154589;
RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted repeats, is conserved in several limited-passage isolates but not in strain 17syn+";
RT J. Virol. 64:1014-1020(1990).
FT REPEAT: 155 172 6 X 3 AA TANDDEM REPEATS OF A-T-P.
SO SEQUENCE 245 AA; 25981 MW; 55ADA50AB18BCA11 CRC64;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M33700; AAA45791.1; -

CC DR Repeat; Late protein.
CC KM DOMAIN 155 172 6 X 3 AA TANDDEM REPEATS OF A-T-P.
SO SEQUENCE 245 AA; 25981 MW; 55ADA50AB18BCA11 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
:111111111
Db 12 RRRPP 17

RESULT 10
ICP3_HSV1 STANDARD; PRT; 248 AA.
ID ICP3_HSV1 P36313;
AC P36313;
DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
GN ICP34.5.
OS Herpes simplex virus (type 1 / strain 17).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McComb D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1";
RT J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=92341080; PubMed=1321882;
RA Dolan A., McKie E., McLean A.R., McGeoch D.J.;
RT "Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17";
RT J. Gen. Virol. 73:971-973(1992).
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X14112; -; NOT_ANNOTATED_CDS.
CC DR PIR: J01682; J01682.
CC KM Repeat.
CC FT DOMAIN 161 175 5 X 3 AA TANDDEM REPEATS OF A-T-P.
SO SEQUENCE 248 AA; 26185 MW; AD344AFC09CABA65 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
:111111111
Db 11 RRRPP 16

RESULT 11
ICP3_HSV1D STANDARD; PRT; 252 AA.
ID ICP3_HSV1D P37318;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
GN ICP34.5.
OS Herpes simplex virus (type 1 / strain CVG-2).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=37106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90156494; PubMed=2154589;
RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted repeats, is conserved in several limited-passage isolates but not in strain 17syn+";
RT J. Virol. 64:1014-1020(1990).

CC -1- FUNCTION: NOT KNOWN: HAS A PROFOUND EFFECT ON THE GROWTH OF THE
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33701; AAA45792.1; -
KW Repeat; Late protein. 6 X 3 AA TANDEM REPEATS OF A-T-P.
FT DOMAIN 162 179
SQ SEQUENCE 252 AA; 26601 MW; A6E6D049FEAA79A CRC64;
Query Match 75.0%; Score 36; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRRPP 7
DB 12 RRRPP 17
RESULT 12
ICP3_HSV1F STANDARD; PRT; 263 AA.
ID ICP3_HSV1F
AC P08353;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE INFECTED CELL PROTEIN ICP34.5 (NEUROVIRULENCE FACTOR ICP34.5).
DE ICP34.5.
OS Herpes simplex virus (type 1 / strain F).
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_Taxid=10304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86115412; PubMed=3003394;
RA Chou J., Roizman B.;
RT "The terminal a sequence of the herpes simplex virus genome contains
RT the promoter of a gene located in the repeat sequences of the L
RT component.";
RL J. Virol. 57:629-637(1986).
RN [2]
RP REVISIONS.
RX MEDLINE=90156494; PubMed=2154589;
RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
RT repeats, is conserved in several limited-passagen isolates but not in
RT strain 17syn+.";
RL J. Virol. 64:1014-1020(1990).
CC -1- FUNCTION: NOT KNOWN: HAS A PROFOUND EFFECT ON THE GROWTH OF THE
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M12240; AAA45794.1; ALT_SEQ.
DR EMBL: M33699; AAA45790.1; -
DR PIR: A27768; WMBE38.

KW Repeat; Late protein.
FT DOMAIN 161 190
SQ SEQUENCE 263 AA; 27533 MW; F5084106A08ACB1 CRC64;
Query Match 75.0%; Score 36; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RRRPP 7
DB 11 RRRPP 16
RESULT 13
C1W4_HUMAN STANDARD; PRT; 393 AA.
ID C1W4_HUMAN
AC Q9NYG8;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 4 (TWIK-RELATED ARACHIDONIC ACID-
DE STIMULATED POTASSIUM CHANNEL PROTEIN) (TRAAK).
GN KCNK4 OR TRAAK.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=11042359;
RA Chapman C.G., Meadows H.J., Godden R.J., Campbell D.A., Duckworth M.,
RA Kelsell R.E., Murdoch P.R., Randall A.D., Renne G.I., Gloger I.S.,
RT "Cloning, localisation and functional expression of a novel human,
RL cerebellum specific, two pore domain potassium channel.";
RL Brain Res. Mol. Brain Res. 82:74-83(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal cortex;
RA Gray A.T.;
RT "Assignment of KCNK4 encoding the human potassium channel TRAAK to
RT chromosome 11.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: VOLTAGE INSENSITIVE, INSTANTANEOUS, OUTWARDLY RECTIFYING
CC POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS REVERSED AT HIGH
CC EXTERNAL K+ CONCENTRATIONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
CC CHANNELS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF248242; AAG31731.1; -
DR EMBL: AF247042; AAF64062.1; ALT_INIT.
DR InterPro: IPR001622; -
DR InterPro: IPR003280; -
DR PRINTS: PR01333; 2PORECHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Potassium transport;
KW Glycoprotein.
FT DOMAIN 1 3
FT TRANSMEM 4 24
FT DOMAIN 89 113
FT TRANSMEM 118 138
FT DOMAIN 140 171
FT TRANSMEM 172 192
CYTOPLASMIC (POTENTIAL).
POTENTIAL.
PORE-FORMING (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
POTENTIAL.


```

FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 638 AA; 72196 MW; F502A510BFE65392 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 638;
Best Local Similarity 75.0%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
   1111: 11
Db 562 RRRPQAPY 569

RESULT 16
SCAB_RAT STANDARD; PRT; 638 AA.
AC P37090; 009183;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE AMILORIDE-SENSITIVE SODIUM CHANNEL BETA-SUBUNIT (EPITHELIAL NA+
DE CHANNEL BETA SUBUNIT) (BETA ENAC) (NONVOLTAGE-GATED SODIUM CHANNEL 1
DE BETA SUBUNIT) (SCNEB) (BETA NACH).
GN SCNN1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

[1]
SEQUENCE FROM N.A.
RA STRAIN=SPRAGUE-DAWLEY; TISSUE=Distal colon epithelium;
RX MEDLINE=94150624; PubMed=8107805;
RA Canessa C.M., Schild L., Buell G., Thorens B., Gautschi I.,
RA Horisberger J.D., Rossier B.C.;
RA "Amiloride-sensitive epithelial Na+ channel is made of three
RA homologous subunits.";
RA Nature 367:463-467(1994).

[2]
SEQUENCE FROM N.A.
RA TISSUE=Kidney;
RX MEDLINE=97191134; PubMed=9039092;
RA Kreutz R., Struck B., Rudatta S., Hubner N., Szpirer J., Szpirer C.,
RA Ganten D., Lindpaintner K.;
RA "Role of the alpha-, beta-, and gamma-subunits of epithelial sodium
RA channel in a model of polygenic hypertension.";
RA Hypertension 29:131-136(1997).

-1- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE-SENSITIVE ION CHANNEL.
INHIBITED BY THE DIURETIC AMILORIDE. MEDIATE THE ELECTRODIFFUSION
OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY)
THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE
REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS.
CC ALSO PLAYS A ROLE IN TASTE PERCEPTION.
-1- SUBUNIT: HETEROTRIMER OF TWO ALPHA, ONE BETA AND ONE GAMMA
SUBUNIT. A DELTA SUBUNIT CAN REPLACE THE ALPHA SUBUNIT.
CC SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-1- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNELS
FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL; X77932; CAA54904.1; -
EMBL; U35174; AAB58457.1; -
DR

```

```

DR EMBL; U35175; AAB58458.1; -
DR InterPro; IPR001873; -
DR Pfam; PF00858; ASC; 1.
DR PRINTS; PR01076; AMINCHANNEL.
DR PROSITE; PS01206; ASC; 1.
KW Ionic channel; Transmembrane; Ion transport; Glycoprotein.
DR CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 50
FT TRANSSEM 51 71
FT DOMAIN 72 71
FT TRANSSEM 72 530
FT DOMAIN 531 551
FT TRANSSEM 552 638
FT DOMAIN 135 135
FT CARBOHYD 141 141
FT CARBOHYD 172 172
FT CARBOHYD 197 197
FT CARBOHYD 205 205
FT CARBOHYD 362 362
FT CARBOHYD 376 376
FT CARBOHYD 396 396
FT CARBOHYD 447 447
FT CARBOHYD 482 482
FT CONFLICT 567 567
FT CONFLICT 586 587
SQ SEQUENCE 638 AA; 71994 MW; 79176A282D5079F1 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 638;
Best Local Similarity 75.0%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRRPPY 8
   1111: 11
Db 562 RRRPQAPY 569

RESULT 17
POLR_RUBVM STANDARD; PRT; 992 AA.
ID POLR_RUBVM
AC P08563;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE STRUCTURAL POLYPEPTIDE [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE
DE GLYCOPROTEIN E1 AND E2].
OS Rubella virus (Strain M33).
OC Viruses; ssRNA Positive-strand viruses, no DNA stage; Togaviridae;
OC Rubivirus.
OX NCBI_TaxID=11043;

[1]
SEQUENCE FROM N.A.
RX MEDLINE=87174825; PubMed=3562245;
RA Clarke D.M., Loo T.W., Hui I., Chong P., Giliam S.;
RA "Nucleotide sequence and in vitro expression of rubella virus 24S
RA subgenomic messenger RNA encoding the structural proteins E1, E2 and
RA C.";
RL Nucleic Acids Res. 15:3041-3057(1987).

-1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ
ACTIVITY.
-1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE
NUCLEOCAPSID.
CC DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY
RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL; X05259; CAA28880.1; -
DR PIR; A27505; GNMV93.
DR

```

KM Polyprotein; Nucleocapsid; Glycoprotein; Transmembrane; Signal.
 FT CHAIN 1 299 NUCLEOCAPSID PROTEIN C.
 FT CARBOHYD 300 559 MEMBRANE GLYCOPROTEIN E2.
 FT SIGNAL 560 580
 FT CHAIN 581 992 MEMBRANE GLYCOPROTEIN E1.
 FT TRANSMEM 277 297 POTENTIAL.
 FT TRANSMEM 515 531 POTENTIAL.
 FT TRANSMEM 533 553 POTENTIAL.
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 757 757 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 992 AA; 106905 MW; D26889C9E78CF1E CRC64;

Query Match
 Best Local Similarity 100.0%; Score 36; DB 1; Length 992;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
 Db 35 RRRPP 40

RESULT 18
 POLS_RUBVR STANDARD; PRT; 1063 AA.
 AC P21480; Q86373; Q86374; Q86375;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE STRUCTURAL POLYPEPTIDE [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE GLYCOPROTEIN E1 AND E2].
 OS Rubella virus (vaccine strain HPV77).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Rubivirus.
 NC NCB1_Taxid-11044;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-90060825; Pubmed-2583526;
 RA Zheng D., Dickens L., Liu T.Y., Nakhasi H.L.;
 FT Nucleotide sequence of the 24S subgenomic messenger RNA of a vaccine strain (HPV77) of rubella virus: comparison with a wild-type strain (M33).
 RT Gene 82:343-349(1989).
 CC -1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ ACTIVITY.
 CC -1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE NUCLEOCAPSID.
 CC -1- DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M30776; AAA47421.1; -;
 DR PIR; J00087; GNMV77.
 KW Polyprotein; Nucleocapsid; Glycoprotein; Transmembrane; Signal.
 FT CHAIN 1 300 NUCLEOCAPSID PROTEIN C.
 FT CARBOHYD 301 562 MEMBRANE GLYCOPROTEIN E2.
 FT SIGNAL 563 582
 FT CHAIN 583 1063 MEMBRANE GLYCOPROTEIN E1.
 FT TRANSMEM 278 300 POTENTIAL.
 FT TRANSMEM 517 535 POTENTIAL.
 FT TRANSMEM 538 555 POTENTIAL.
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1063 AA; 114713 MW; 18E77B28E89A4667 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 36; DB 1; Length 1063;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
 Db 35 RRRPP 40

RESULT 19
 POLS_RUBVR STANDARD; PRT; 1063 AA.
 AC P19725; Q86370; Q86371; Q86372;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE STRUCTURAL POLYPEPTIDE [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE GLYCOPROTEIN E1 AND E2].
 OS Rubella virus (vaccine strain RA27/3).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Rubivirus.
 NC NCB1_Taxid-11044;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-89296505; Pubmed-2740235;
 RA Nakhasi H.L., Thomas D., Zheng D., Liu T.Y.;
 FT Nucleotide sequence of capsid, E2 and E1 protein genes of Rubella virus vaccine strain RA27/3.
 RT Nucleic Acids Res. 17:4393-4394(1989).
 CC -1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ ACTIVITY.
 CC -1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE NUCLEOCAPSID.
 CC -1- DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X14871; CAA33016.1; -;
 DR PIR; S04800; GNMVRA.
 KW Polyprotein; Nucleocapsid; Glycoprotein; Transmembrane; Signal.
 FT CHAIN 1 300 NUCLEOCAPSID PROTEIN C.
 FT CHAIN 301 562 MEMBRANE GLYCOPROTEIN E2.
 FT SIGNAL 563 582 MEMBRANE GLYCOPROTEIN E1.
 FT CHAIN 583 1063
 FT TRANSMEM 278 300 POTENTIAL.
 FT TRANSMEM 517 535 POTENTIAL.
 FT TRANSMEM 538 555 POTENTIAL.
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1063 AA; 114762 MW; 7A1585380C967EFC CRC64;

Query Match
 Best Local Similarity 100.0%; Score 36; DB 1; Length 1063;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 |||||
 Db 35 RRRPP 40

RESULT 20
 POLS_RUBVT STANDARD; PRT: 1063 AA.

AC P07566;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE STRUCTURAL POLYPROTEIN [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE GLYCOPROTEIN E1 AND E2].
 DE Rubella virus (strain Therien).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Rubivirus.
 OC NCBI_TaxID=11045.

RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281585; PubMed=2353453;
 RA Dominguez G., Wang C.Y., Frey T.K.;
 RT "Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution.";
 RT Virology 177:225-238(1990).
 RL

RP [2]
 RP SEQUENCE OF 1-621 FROM N.A.
 RX MEDLINE=8826020; PubMed=2836271;
 RA Frey T.K., Marr L.D.;
 RT "Sequence of the region coding for virion proteins C and E2 and the carboxy terminus of the nonstructural proteins of rubella virus: comparison with alphaviruses.";
 RT Gene 62:85-99(1988).
 RL

RP [3]
 RP SEQUENCE OF 551-1063 FROM N.A.
 RX MEDLINE=86317717; PubMed=3755848;
 RA Frey T.K., Marr L.D., Hemphill M.L., Dominguez G.;
 RT "Molecular cloning and sequencing of the region of the rubella virus genome coding for glycoprotein E1.";
 RT Virology 154:228-232(1986).
 RL

RP [4]
 RP SEQUENCE OF 1-319 FROM N.A.
 RX MEDLINE=88171449; PubMed=3351478;
 RA Takkinen K., Vidgren G., Ekstrand J., Hellman U., Kalkkinen N., Wernstedt C., Petersson R.F.;
 RT "Nucleotide sequence of the rubella virus capsid protein gene reveals an unusually high G/C content.";
 RT J. Gen. Virol. 69:603-612(1988).
 RL

RP [5]
 RP SEQUENCE OF 278-1063 FROM N.A.
 RX MEDLINE=88009948; PubMed=3655744;
 RA Vidgren G., Takkinen K., Kalkkinen N., Kaaerilaenen L., Petersson R.F.;
 RT "Nucleotide sequence of the genes coding for the membrane glycoproteins E1 and E2 of rubella virus.";
 RT J. Gen. Virol. 68:2347-2357(1987).
 RL

RP -1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ ACTIVITY.
 RP -1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE NUCLEOCAPSID.
 RP -1- DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M15240; AAA88529.1; -;
 DR EMBL; D00242; BAA00172.1; -;
 DR EMBL; D00156; BAA28178.1; -;
 DR PIR; A29822; GNMVRA.
 DR PIR; A24309; A24309.
 KW Polyprotein; Nucleocapsid; Glycoprotein; Transmembrane; Signal.
 FT CHAIN 1 300 MEMBRANE GLYCOPROTEIN E1.
 FT CHAIN 1 300 NUCLEOCAPSID PROTEIN C.
 FT SIGNAL 562 582 MEMBRANE GLYCOPROTEIN E2.
 FT CHAIN 583 1063
 FT TRANSMEM 278 300
 FT TRANSMEM 517 535
 FT TRANSMEM 538 555
 FT CARBOHYD 353 353
 FT CARBOHYD 371 371
 FT CARBOHYD 410 410
 FT CARBOHYD 429 429
 FT CARBOHYD 658 658
 FT CARBOHYD 759 759
 FT CARBOHYD 791 791
 FT CONFLICT 87 87
 FT CONFLICT 163 163
 FT CONFLICT 319 319
 FT CONFLICT 395 395
 FT CONFLICT 545 545
 FT CONFLICT 959 959
 FT CONFLICT 991 991
 FT CONFLICT 1037 1037
 SO SEQUENCE 1063 AA; 114678 MW; F39B475ACAI5C7D1 CRC64;

Query Match 75.0%; Score 36; DB 1; Length 1063;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 |||||
 Db 35 RRRPP 40

RESULT 21
 SSS2_SCYCA STANDARD; PRT: 80 AA.

ID SSS2_SCYCA
 AC P11020;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE SPERMATID-SPECIFIC PROTEIN S2 (BASIC NUCLEAR PROTEIN S2).
 OS Scyliorhinus canicula (Spotted dogfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;
 OC Scyliorhinidae; Scyliorhinus.
 OX NCBI_TaxID=7830;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=89170733; PubMed=2924768;
 RA Chauviere M., Martinge A., Briand G., Sautiere P., Chevaillier P.;
 RT "Nuclear basic protein transition during sperm differentiation. Primary structure of the spermatid-specific protein S2 from the dog-fish Scyliorhinus caniculus.";
 RT Eur. J. Biochem. 180:329-335(1989).
 RL

RP -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION. HISTONES ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES REPLACED BY PROTAMINES IN LATE SPERMATIDS.
 RP -1- SUBCELLULAR LOCATION: NUCLEAR.
 RP -1- MISCELLANEOUS: N-TERMINAL HALF IS HIGHLY BASIC, WHILE C-TERMINAL PART IS ACID.
 RP -1- SIMILARITY: WITH PROTEIN S1 (70% STRUCTURAL SIMILARITY).
 DR PIR; S03560; S03560.
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding; Nuclear protein.
 SO SEQUENCE 80 AA; 9733 MW; 593E18EF3452409 CRC64;

RT Location.":
J. Biochem. 111:123-128(1992).
- FUNCTION: ACTIVATION OF LONG-CHAIN FATTY ACIDS FOR BOTH SYNTHESIS
OF CELLULAR LIPIDS, AND DEGRADATION VIA BETA-OXIDATION.
- CATALYTIC ACTIVITY: ATP + A LONG-CHAIN CARBOXYLIC ACID + COA = AMP
+ PYROPHOSPHATE + AN ACYL-COA.
- COFACTOR: REQUIRES MAGNESIUM.
- SUBCELLULAR LOCATION: MICROSOMES, OUTER MITOCHONDRIAL MEMBRANE
AND PEROXISOMAL MEMBRANE.
- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: D10040; BAA00931.1; -
DR PIR: JX0202; JX0202.
DR HSP: P08659; ILCT.
DR MIM: 152426; -
DR InterPro: IPR000873; -
DR Pfam: PF00501; AMP-binding; 1.
DR PRINTS: PR00154; AMP-BINDING.
DR PROSITE: PS00455; AMP-BINDING; 1.
DR Ligase: Fatty acid metabolism; Magnesium; Multigene family.
KW SEQUENCE 698 AA; 77943 MW; FD669453589D362 CRC64;
SQ

Query Match 72.9%; Score 35; DB 1; Length 698;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRP 7
111111
DB 259 RRRPRP 265

RESULT 24
LCFA_HUMAN STANDARD; PRT; 699 AA.
AC PA1215;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LONG-CHAIN-FATTY-ACID-COA LIGASE 1 (EC 6.2.1.3) (LONG-CHAIN ACYL-COA
SYNTHETASE 1) (LACS 1) (PALMITOYL-COA LIGASE).
DE FACIL OR LACS1 OR LACS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96147073; PubMed=8564017;
RA Ghosh B., Batdosa E., Singh I.;
RT "Molecular cloning and sequencing of human palmitoyl-CoA ligase and
its tissue specific expression.";
RL Mol. Cell. Biochem. 151:77-81(1995).
CC -1- FUNCTION: ACTIVATION OF LONG-CHAIN FATTY ACIDS FOR BOTH SYNTHESIS
OF CELLULAR LIPIDS, AND DEGRADATION VIA BETA-OXIDATION.
CC -1- CATALYTIC ACTIVITY: ATP + A LONG-CHAIN CARBOXYLIC ACID + COA = AMP
+ PYROPHOSPHATE + AN ACYL-COA.
CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: MICROSOMES, OUTER MITOCHONDRIAL MEMBRANE
AND PEROXISOMAL MEMBRANE.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, HEART, SKELETAL
MUSCLE AND KIDNEY, AND TO A LESSER EXTENT IN BRAIN, LUNG, PLACENTA
AND PANCREAS.
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT

CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L09229; AAB00959.1; -
DR HSP: P08659; ILCT.
DR MIM: 152425; -
DR InterPro: IPR000873; -
DR Pfam: PF00501; AMP-binding; 1.
DR PRINTS: PR00154; AMP-BINDING.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW Ligase: Fatty acid metabolism; Magnesium; Multigene family.
SQ SEQUENCE 699 AA; 78347 MW; 9EDF6FD8AF20F07B CRC64;
SQ

Query Match 72.9%; Score 35; DB 1; Length 699;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPRP 7
111111
DB 259 RRRPRP 265

RESULT 25
PKSL_ASPPA STANDARD; PRT; 2109 AA.
AC PKSL_ASPPA
ID Q12053;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE AFLATOXIN BIOSYNTHESIS POLYKETIDE SYNTHASE (PKS).
GN PKSL1.
OS Aspergillus parasiticus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5067;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NRRL 2999;
RX MEDLINE=96042102; PubMed=7592391;
RA Peng G.H., Leonard T.J.;
RT Characterization of the polyketide synthase gene (pksl1) required
for aflatoxin biosynthesis in Aspergillus parasiticus.";
RL J. Bacteriol. 177:6246-6254(1995).
CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
AFLATOXIN FROM HEXANOYL COA AND SEVEN MALONATES.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE
(POTENTIAL).
CC -1- PATHWAY: FIRST STEP IN AFLATOXIN BIOSYNTHESIS.
CC -1- SIMILARITY: WITH BOTH EUKARYOTIC AND PROKARYOTIC POLYKETIDE
SYNTHASES AND VERTEBRATE FATTY ACID SYNTHASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L42766; AAC1675.1; -
DR EMBL: L42765; AAC1674.1; -
DR InterPro: IPR000255; -
DR PRINTS: PR000794; -
DR InterPro: IPR001031; -
DR InterPro: IPR001227; -


```

DR Pfam: PF00698; Acyl_transf; 1.
DR Pfam: PF00975; Thioesterase; 1.
DR Pfam: PF0109; ketoacyl-synt; 1.
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS50075; ACP_DOMAIN; 1.
KW Transferrase; Acyltransferase; Phosphopantetheine;
KW Multifunctional enzyme.
FT DOMAIN 374 805 BETA-KETOACYL SYNTHASE.
FT DOMAIN ? ? ACYL/MALONYL TRANSFERASES.
FT DOMAIN 1714 1785 ACYL CARRIER (ACP).
FT DOMAIN ? 2109 THIOESTERASE.
FT ACT_SITE 543 543 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 993 993 ACYL/MALONYL TRANSFERASES (BY SIMILARITY).
FT BINDING 1746 1746 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT_SITE 1937 1937 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 2109 AA; 230715 MW; CB701372A16D8551 CRC64;

Query Match
Best Local Similarity 72.9%; Score 35; DB 1; Length 2109;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPPY 8
   |||||
Db 1924 RRRPRGPY 1931

RESULT 26
MK1_PALPR STANDARD; PRT; 15 AA.
AC P80408;
DR 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DR 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALINKOWIN 1.
OS Palomera prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Palomera.
OX NCBI_TaxID=55431;
RN [1]
RS SEQUENCE.
RC TISSUE=Hemolymph;
RA Chenyuh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The Inducible antibacterial peptides of the hemipteran insect
RT Palomera prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin."
RL J. Insect Physiol. 42:81-89(1996).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match
Best Local Similarity 70.8%; Score 34; DB 1; Length 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPP 7
   |||||
Db 7 RRRPRPP 13

RESULT 27
XG_HUMAN STANDARD; PRT; 180 AA.
ID XG_HUMAN
AC P55808;
DR 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DR 01-OCT-2000 (Rel. 40, Last annotation update)

```

```

DE XG GLYCOPROTEIN PRECURSOR (PROTEIN PBDX).
GN XG OR PBDX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=94332149; PubMed=8054981;
RA Ellis N.A., Ye T.Z., Patton S., German J., Goodfellow P.N.,
RA Weller P.;
RT "Cloning of PBDX, an MIC2-related gene that spans the pseudoautosomal
RT boundary on chromosome Xp."
RL Genet. 6:394-400(1994).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -1- PTM: O-GLYCOSYLATED (PROBABLE).
CC -1- POLYMORPHISM: XG IS RESPONSIBLE FOR THE XG BLOOD GROUP SYSTEM.
CC -1- SIMILARITY: TO PROTEIN MIC2/CD99.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X96421; -; NOT_ANNOTATED_CDS.
DR MIM: 314700; -
KW Transmembrane; Glycoprotein; Blood group antigen; Signal.
FT SIGNAL 1 21
FT CHAIN 22 180 XG GLYCOPROTEIN.
FT TRANSMEM 143 163 POTENTIAL.
SQ SEQUENCE 180 AA; 19723 MW; DADAA9E6859C4530 CRC64;

Query Match
Best Local Similarity 70.8%; Score 34; DB 1; Length 180;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRPP 7
   |||||
Db 100 RRRPRPP 106

RESULT 28
Y215_ADE02 STANDARD; PRT; 215 AA.
ID Y215_ADE02
AC P03291;
DR 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DR 01-APR-1990 (Rel. 14, Last annotation update)
DE HYPOTHETICAL PROTEIN F-215.
OS Human adenovirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515;
RN [1]
RS SEQUENCE FROM N.A.
RP MEDLINE=83056843; PubMed=7142161;
RX Gingeras T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
RT "Nucleotide sequences from the adenovirus-2 genome."
RL J. Biol. Chem. 257:13475-13491(1982).
RN [2]
RS SEQUENCE FROM N.A.
RP MEDLINE=83056844; PubMed=7142162;
RX Alestroom P., Akusjaervi G., Pettersson M., Pettersson U.;
RT "DNA sequence analysis of the region encoding the terminal protein
RT and the hypothetical N-gene product of adenovirus type 2."
RL J. Biol. Chem. 257:13492-13498(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: J01917; -- NOT_ANNOTATED_CDS.
CC PIR: A03863; A03863.
CC Hypothetical protein.
CC
SQ SEQUENCE 215 AA; 22787 MW; 94FA33BAAB7AF137 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
| | | | |
Db 135 RRVSPPY 142

RESULT 29
RL_HSV2H STANDARD; PRT; 261 AA.
AC P28283;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NEUROVIRULENCE FACTOR (ICP34.5).
RN 1.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RN J. Gen. Virol. 72:3057-3075(1991).
RL 12;
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: D10471; BAA23428.1; --
CC EMBL: Z86099; CAB06759.1; --
CC EMBL: Z86099; CAB06706.1; --
CC PIR: J01502; WMBEXE.
CC Repeat.
KW DOMAIN 3 12 2 X 5 AA TANDEM REPEATS OF R-R-R-G-P.
FT REPEAT 3 7
FT REPEAT 8 12
FT REPEAT 16 31 2 X 8 AA TANDEM REPEATS OF P-R-P-G-A-P-A-
FT DOMAIN 16 31 V.
FT REPEAT 16 23
FT REPEAT 24 31
SQ SEQUENCE 261 AA; 27908 MW; 4BBD13AF3D906D71 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 69;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 6
| | | | |
Db 13 RRRPRP 18

RESULT 30
RS31_ARATH STANDARD; PRT; 264 AA.
ID RS31_ARATH
AC P92964;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ARGININE/SERINE-RICH SPLICING FACTOR RSP31.
GN RSP31
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=97143875; PubMed=8989882;
RA Lopato S., Waigmann E., Barta A.;
RT "Characterization of a novel arginine/serine-rich splicing factor in
RT Arabidopsis."
RL Plant Cell 6:2255-2264(1996).
CC -1- FUNCTION: REQUIRED FOR CONSTITUTIVE AND ALTERNATIVE PRE-mRNA
CC SPLICING.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ROOTS AND FLOWERS. A
CC PRESUMABLY LONGER ALTERNATIVE SPLICED FORM IS FOUND IN LEAVES,
CC STEMS AND FLOWERS.
CC -1- SIMILARITY: BELONGS 2 RNA RECOGNITION MOTIFS (RRM).
CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: X99435; CAA67798.1; --
CC InterPro: IPR000504; --
CC Pfam: PF00076; Rrm; 2.
DR PROSITE: PS0102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
CC Nucleic acid binding protein; RNA-binding; Alternative splicing; Repeat.
KW DOMAIN 2 74
FT REPEAT 2 74
FT REPEAT 93 164
FT DOMAIN 172 264
SQ SEQUENCE 264 AA; 31126 MW; 64C7988166FF719 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 264;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRPRPY 8
| | | | |
Db 187 RRRSPDY 194

RESULT 31
HME1_CHICK STANDARD; PRT; 333 AA.
ID HME1_CHICK
AC Q05916;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOMEBOX PROTEIN ENGRAILED-1 (Gg-EN-1).
 EN1 OR EN-1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 RN NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93185339; PubMed=1363401;
 RA Logan C., Hanks M.C., Noble-Topham S., Nallanathan D.,
 RA Provart N.J., Joyner A.L.;
 RT "Cloning and sequence comparison of the mouse, human, and chicken
 RT engrailed genes reveal potential functional domains and regulatory
 RT regions.";
 RL Dev. Genet. 13:345-358(1992).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE ENGRAILED FAMILY OF HOMEBOX PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12694; AAA53435.1; -;
 DR EMBL: L12695; AAA53436.1; -;
 DR HSP: P02836; IHDD.
 DR TRANSFAC: T02014; -;
 DR InterPro: IPR000747; -;
 DR InterPro: IPR001356; -;
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR PRINTS: PR00026; ENGRAILED.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 DR PROSITE: PS00033; ENGRAILED; 1.
 KM Homeobox: DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 22 29 POLY-GLY.
 FT DOMAIN 36 65 PRO-RICH.
 FT DOMAIN 59 65 POLY-PRO.
 FT DOMAIN 97 104 POLY-GLY.
 FT DNA_BIND 244 303 HOMEBOX.
 SQ SEQUENCE 333 AA; 34515 MW; DQF1B1F917EIFBAD CRC64;
 Query Match 70.8%; Score 34; DB 1; Length 333;
 Best Local Similarity 85.7%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RRRPRP 7
 DB 56 RRRPRP 62
 RESULT 32
 DBPA_RAT STANDARD; PRT; 361 AA.
 ID DBPA_RAT
 AC 062764; Q63748;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA-BINDING PROTEIN A (COLD SHOCK DOMAIN PROTEIN A) (MUSCLE Y-BOX
 DE PROTEIN YB2) (Y-BOX BINDING PROTEIN-A) (RYB-A).
 GN CSDA OR DBPA OR YB2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Goldaman D., Gao J., Burmeister M., Sapru M.;
 RT "Characterization of muscle Y-box proteins that bind the nACHR
 RT delta subunit promoter.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; Tissue=Liver;
 RX MEDLINE=94301785; PubMed=8029009;
 RA Ito K., Tsutsuni K., Kuzumaki T., Gomez P.F., Otsu K., Ishikawa K.;
 RT "A novel growth-inducible gene that encodes a protein with a
 RT conserved cold-shock domain.";
 RL Nucleic Acids Res. 22:2036-2041(1994).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN THE SKELETAL MUSCLE, SPLEEN, AND
 CC FETAL LIVER.
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U22893; AAB60520.1; -;
 DR EMBL: D28557; BAA05907.1; -;
 DR InterPro: IPR002059; -;
 DR Pfam: PF00313; CSD; 1.
 DR PRINTS: PR00050; COLDSHOCK.
 DR PROSITE: PS00352; COLD SHOCK; 1.
 KM Transcription regulation; Repressor; DNA-binding; Nuclear protein;
 KW Alternative splicing.
 FT DOMAIN 85 149 CSD.
 FT VARSPPLIC 184 252 MISSING (IN ISOFORM 2).
 FT CONFLICT 14 14 L->H (IN REF. 2).
 FT CONFLICT 52 74 SPGDADGPPAPASSAPAGSADA -> APARASPARPGLI
 FT SPRGKRG (IN REF. 2).
 FT PR->HY (IN REF. 2).
 SQ SEQUENCE 120 121 PR->HY (IN REF. 2).
 Query Match 70.8%; Score 34; DB 1; Length 361;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRRPRP 6
 DB 325 RRRPRP 330
 RESULT 33
 RRG2_HUMAN STANDARD; PRT; 443 AA.
 ID RRG2_HUMAN
 AC P22932;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-2 (RAR-GAMMA-2).
 GN RARG OR NR1B3.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-63 FROM N.A.
 RX MEDLINE=90207264; PubMed=2157210.
 RA Kastner P., Krust A., Mendelsohn C., Garnier J.M., Zelent A.,
 RA Leroy P., Staub A., Chambon P.;

RT "Murine isoforms of retinoic acid receptor gamma with specific
 RT patterns of expression."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2700-2704(1990).
 RN [2]
 RP SEQUENCE OF 51-443 FROM N.A.
 RA MEDLINE=89315787; PubMed=2546152;
 RX Krust A., Kastner P., Petkovich M., Zelent A., Chambon P.;
 RA "A third human retinoic acid receptor, hRAR-gamma,"
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5310-5314(1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 167-412.
 RA MEDLINE=96107307; PubMed=7501014;
 RX Renaud J.-P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H.,
 RA Moras D.;
 RA "Crystal structure of the RAR-gamma ligand-binding domain bound to
 RL all-trans retinoic acid,"
 RL Nature 378:681-689(1995).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 167-412.
 RA MEDLINE=98162555; PubMed=9501913;
 RX Klaholz B.P., Renaud J.-P., Mitschler A., Zusi C., Chambon P.,
 RA Gronemeyer H., Moras D.;
 RA "Conformational adaptation of agonists to the human nuclear receptor
 RT RAR gamma,"
 RL Nat. Struct. Biol. 5:199-202(1998).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 CC HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 CC A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CC CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: RAR-GAMMA-1 AND -2 ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY DIFFER IN THEIR
 CC N-TERMINAL SECTION.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M24857; NOT ANNOTATED_CDS.
 DR EMBL; M32074; AAA60253.1;
 DR PDB; 3LBD; 02-MAR-99.
 DR PDB; 4LBD; 02-MAR-99.
 DR TRASNPA; T00720;
 DR MIM; 180190;
 DR InterPro; IPR000536;
 DR InterPro; IPR001628;
 DR InterPro; IPR001723;
 DR InterPro; IPR003078;
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PRO00047; STEROIDFINGER.
 DR PRINTS; PRO0398; STRDHOMER.
 DR PRINTS; PRO1292; RETNOTACIDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Multigene family; Alternative splicing; 3D-structure.
 FT DOMAIN 1 78
 FT DNAS_BIND 79 144 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 79 99 C4-TYPE.
 FT ZN_FING 115 139 C4-TYPE.
 FT DOMAIN 145 190 HINGE.
 FT DOMAIN 191 410 LIGAND-BINDING.
 SEQUENCE 443 AA; 49307 MW; 4D709194F511B86 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 443;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 RRRRPPY 8
 Db 356 RRRRPPY 363
 RESULT 34
 RS01_HUMAN STANDARD; PRT; 454 AA.
 ID RS01_HUMAN
 AC P13631;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE RETINOIC ACID RECEPTOR GAMMA-1 (RAR-GAMMA-1).
 GN RARG OR NR1B3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89315787; PubMed=2546152;
 RA Krust A., Kastner P., Petkovich M., Zelent A., Chambon P.;
 RT "A third human retinoic acid receptor, hRAR-gamma,"
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5310-5314(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91042616; PubMed=2172793;
 RA Ishikawa T., Umehono K., Mangelsdorf D.J., Aburatani H., Stanger B.Z.,
 RA Shibasaki Y., Inawari M., Evans R.M., Takaku F.;
 RT "A functional retinoic acid receptor encoded by the gene on human
 RT chromosome 12,"
 RL Mol. Endocrinol. 4:837-844(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91187677; PubMed=1849262;
 RA Lehmann J.M., Hoffmann B., Pfahl M.;
 RT "Genomic organization of the retinoic acid receptor gamma gene,"
 RL Nucleic Acids Res. 19:573-578(1991).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 178-423.
 RX MEDLINE=96107307; PubMed=7501014;
 RA Renaud J.-P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H.,
 RA Moras D.;
 RA "Crystal structure of the RAR-gamma ligand-binding domain bound to
 RT all-trans retinoic acid,"
 RL Nature 378:681-689(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 178-423.
 RX MEDLINE=98162555; PubMed=9501913;
 RA Klaholz B.P., Renaud J.-P., Mitschler A., Zusi C., Chambon P.,
 RA Gronemeyer H., Moras D.;
 RA "Conformational adaptation of agonists to the human nuclear receptor
 RT RAR gamma,"
 RL Nat. Struct. Biol. 5:199-202(1998).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 CC HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 CC A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CC CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: RAR-GAMMA-1 AND -2 ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY DIFFER IN THEIR
 CC N-TERMINAL SECTION.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M24857; AAA52692.1; -
 DR EMBL: M38258; AAA60254.1; -
 DR EMBL: M57707; AAA63254.1; -
 DR PIR: A33903; A33903.
 DR PIR: A35573; A35573.
 DR PIR: S26848; S26848.
 DR PDB: 3LBD; 02-MAR-99.
 DR PDB: 4LBD; 02-MAR-99.
 DR TRANSFAC: T00720; -
 DR TRANSFAC: T01330; -
 DR MIM: 180190; -
 DR InterPro: IPR000536; -
 DR InterPro: IPR001628; -
 DR InterPro: IPR001723; -
 DR InterPro: IPR003078; -
 DR Pfam: PF00104; hormone_rec; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDEFINGER.
 DR PRINTS: PR00398; STROHOMONER.
 DR PRINTS: PR01292; RETNOICACIDR.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 DR Zinc-finger; Multigene family; Alternative splicing; 3D-structure.
 FM DOMAIN 1 89 MODULATING.
 FT DNA_BIND 90 155 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 90 110 C4-TYPE.
 FT ZN_FING 126 150 C4-TYPE.
 FT DOMAIN 156 201 HINGE.
 FT DOMAIN 202 421 LIGAND-BINDING.
 SQ SEQUENCE 454 AA; 50341 MW; 1EE27B22772DAFDCRC64;
 Query Match 70.8%; Score 34; DB 1; Length 454;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RRRRPY 8
 DB 367 RRRPSQPY 374
 C384_ARATH 35
 ID CP84_ARATH STANDARD; PRT; 520 AA.
 AC 042600;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 GN CYTOCHROME P450 8A1 (FERULATE-5-HYDROXYLASE) (EC 1.14.-.-) (FSH).
 OR CYP84A1 OR FAH1 OR A14G36220 OR F23E13.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 CC [1]
 CC SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA MEDLINE=96293440; PubMed=8692910;
 RA Meyer R., Cusumano J.C., Somerville C.R., Chapple C.C.S.;
 RT "Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new
 RT family of cytochrome P450-dependent monooxygenases.",
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6869-6874(1996).
 RM [2]
 RM SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;

RA Mayer R.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansgore W., Brandt P., Grivell L., Rieger M.,
 RA Weinsalgattner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reihert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Delcroix E.,
 RA Weltzenegger T., Botne G., Rampsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koeter P.,
 RA Beier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysaert C., Gladen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McRay K., Mayes R.,
 RA Pettelt A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herl A.,
 RA Neumann S., Argitrou A., Vitale D., Ligotti R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cheifor F., Cooke R., Berger C., Monfort A., Cascuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Hebermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Clout J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.",
 RL Nature 402:769-777(1999).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV. LANDSBERG ERECTA;
 RC MEDLINE=99097044; PubMed=9880351;
 RX Ruegger M., Meyer R., Cusumano J.C., Chapple C.;
 RT "The regulation of ferulate-5-hydroxylase expression in Arabidopsis in
 RT the context of sinapate ester biosynthesis.",
 RL Plant Physiol. 119:101-110(1999).
 CC -1- PATHWAY: GENERAL PHENYLPROPANOID PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

KM Oxioreductase: Monooxygenase; Membrane; Heme-
FT BINDING 458 HEME (BY SIMILARITY).
SQ SEQUENCE 520 AA; 58720 MW; E812779AF5BF01BC CRC64;

Query Match 70.8%; Score 34; DB 1; Length 520;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPRP 8
1 1 1 1 1 1
Db 34 RRRPRP 40

RESULT 36

GAG_SIVAG1 STANDARD; PRT; 520 AA.

AC P27972;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE GAG POLYPROTEIN [CONTAINS: CORE PROTEINS P17, P24, AND P15].

GN GAG.

OS Simian immunodeficiency virus (AGM15 isolate) (SIV-AGM).

OC Viruses: Retrovird viruses; Retroviridae; Lentivirus.

CC NCBI_TaxID=11727;

CC [1]

CC SEQUENCE FROM N.A.

CC MEDLINE=90156504; PubMed=2304139;

CC Johnson P.R., Fomsgard A., Allan J., Gravel M., London W.T.,

CC Olmstead R.A., Hirsch V.M.;

CC "Simian immunodeficiency viruses from African green monkeys display

CC unusual genetic diversity.";

CC J. Virol. 64:1086-1092(1990).

CC -1- MISCELLANEOUS: THE 155 ISOLATE IS FROM A MONKEY IMPORTED FROM

CC KENYA.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M29975; AAA91905.1; -

CC HSSP: P05888; IAAF.

CC InterPro: IPR000071; -

CC InterPro: IPR001878; -

CC Pfam: PF00540; gag_P17; 1.

CC Pfam: PF00607; gag_P24; 1.

CC Pfam: PF00098; zf-CCHC; 2.

CC PRINTS: PR00234; HYIYMATRFX.

CC PRINTS: PR00339; C2HCZNFINGER.

CC AIDS: Core protein; Polyprotein.

CC CHAIN 1 141

CC FT CHAIN 142 520

CC SEQUENCE 520 AA; 57735 MW; 2FELB7D0EA84D14 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 520;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRP 7
1 1 1 1 1 1
Db 391 RRRPRP 397

RESULT 37

GAG_SIVAG STANDARD; PRT; 521 AA.

AC P27972;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE GAG POLYPROTEIN [CONTAINS: CORE PROTEINS P17, P24, AND P15].

GN GAG.

OS Simian immunodeficiency virus (AGM3 isolate) (SIV-AGM).

OC Viruses: Retrovird viruses; Retroviridae; Lentivirus.

CC NCBI_TaxID=11730;

CC [1]

CC SEQUENCE FROM N.A.

CC MEDLINE=90232731; PubMed=2156689;

CC Baier M., Garber C., Mueller C., Cichutek K., Kuth R.,

CC "Complete nucleotide sequence of a simian immunodeficiency virus from

CC African green monkeys: a novel type of intragroup divergence.";

CC Virolgy 176:216-221(1990).

CC -1- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M30931; AAA91913.1; -

CC HSSP: P05888; IAAF.

CC InterPro: IPR000071; -

CC InterPro: IPR001878; -

CC Pfam: PF00540; gag_P17; 1.

CC Pfam: PF00607; gag_P24; 1.

CC Pfam: PF00098; zf-CCHC; 2.

CC PRINTS: PR00234; HYIYMATRFX.

CC PRINTS: PR00339; C2HCZNFINGER.

CC AIDS: Core protein; Polyprotein.

CC CHAIN 1 145

CC FT CHAIN 146 521

CC SEQUENCE 521 AA; 58409 MW; 1F11BD2F2EDF4F5 CRC64;

Query Match 70.8%; Score 34; DB 1; Length 521;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRP 7
1 1 1 1 1 1
Db 395 RRRPRP 401

RESULT 38

GAG_SIVAG STANDARD; PRT; 587 AA.

AC P53933;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE HYPOTHETICAL 66.1 KDA PROTEIN IN RPS7B-YF753 INTERGENIC REGION.

GN YNL094W OR N2219.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

CC NCBI_TaxID=4932;

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN=5288C / FY1679;

CC MEDLINE=96367601; PubMed=8771715;

CC Garcia-Cantalejo J.M., Boskovic J., Jimenez A.;

CC "Sequence analysis of a 14.2 Kb fragment of Saccharomyces cerevisiae

CC chromosome XIV that includes the ypt53, trnAlu and gsr m2 genes and

CC four new open reading frames.";

CC Yeast 12:599-608(1996).

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X85811; CAA59823.1; ALT_INIT.
DR EMBL: 271370; CAA95970.1; -.
DR SGD: S0005038; YNL094W.
DR Hypothetical protein.
SO SEQUENCE 587 AA; 66134 MW; B85C525548BA34BC CRC64;

Query Match
Best Local Similarity 70.8%; Score 34; DB 1; Length 587;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 RRRRPP 7
    ||| |
    466 RRRRPP 472

RESULT 39
YGFY_ECOLI STANDARD; PRT; 639 AA.
AC Q46820;
DE 01-NOV-1997 (Rel. 35, Created)
DE 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE HYPOTHETICAL 69.1 KDA PROTEIN IN KDOI-LYSS INTERGENIC REGION.
YGFY.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;

[1]
SEQUENCE FROM N.A.
RA STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
CC Science 277:1453-1474(1997).
CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (POTENTIAL).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -1- SIMILARITY: STRONG, TO E.COLI GLUTAMATE SYNTHASE [NADPH] SMALL
CC CHAIN (GLND) AND TO E.COLI AEGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28375; AAA83068.1; ALT_INIT.
DR EMBL: AE000372; AAC75925.1; ALT_INIT.
DR HSSP: P00198; 1FDN.
DR Ecogene; EG13070; YGFY.
DR InterPro: IPR000759; -.
DR InterPro: IPR001450; -.
DR Pfam: PF00037; fer4; 1.
DR PRINTS: PRO0033; 4FE4SFDOXIN.
DR PRINTS: PRO0419; ADXRPTASE.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
DR Hypothetical protein; Oxidoreductase; Iron-sulfur; 4fe-4s.
KW

```

```

FT METAL 87 87 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 90 90 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 93 93 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 97 97 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 258 258 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 262 262 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 268 268 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 272 272 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
SO SEQUENCE 639 AA; 69088 MW; 3B2B96621B35655C CRC64;

Query Match
Best Local Similarity 70.8%; Score 34; DB 1; Length 639;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRRPP 6
    ||| |
    541 RRRRPP 546

RESULT 40
BRD2_HUMAN STANDARD; PRT; 801 AA.
ID BRD2_HUMAN
AC P25440;
DE 01-MAY-1992 (Rel. 22, Created)
DE 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE BROMODOMAIN-CONTAINING PROTEIN 2 (RING3 PROTEIN).
GN BRD2 OR RING3 OR KIAA9001.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

[1]
SEQUENCE FROM N.A.
RA TISSUE-T-cell;
RC MEDLINE-92329974; PubMed-1352711;
RX Beck S., Hanson I., Kelly A., Pappin D.J.C., Trowsdale J.;
RT "A homologue of the Drosophila female sterile homeotic (fsh) gene in
RT the class II region of the human MHC."
RL DNA Seq. 2:203-210(1992).

[2]
REVISIONS TO N-TERMINUS.
RX MEDLINE-96376536; PubMed-8781126;
RA Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
RT "Phylogeny and structure of the RING3 gene."
RL Immunogenetics 44:391-396(1996).

[3]
SEQUENCE FROM N.A.
RA TISSUE-Bone marrow;
RC Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki T., Ishikawa K.-I., Tabata S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -1- SIMILARITY: CONTAINS 1 ET DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62083; CAA4396.1; -.
DR EMBL: M80613; AAA68890.1; ALT_INIT.
DR EMBL: DA2040; BAA07641.1; -.
DR PIR: S18860; S18860.
DR PIR: S40781; S40781.
DR MIM: 601540; -.
DR InterPro: IPR001487; -.
DR Pfam: PF00439; bromodomain; 2.

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF151141; AAD42089.1; -
 DR MGD: MGI:1341105; Rplh.
 DR PROSITE: PS50309; DC; 2.
 KW Vision; Retinal protein; Repeat.
 FT NON_TER 1 1
 FT DOMAIN <1 51 DOUBLECORTIN.
 FT DOMAIN 91 170 DOUBLECORTIN.
 FT DOMAIN 589 594 POLY-LYS.
 FT NON_TER 936 936
 SQ SEQUENCE 936 AA; 104868 MW; 3B1D4B34F0FD69D8 CRC64;

Query Match
 Best Local Similarity 70.8%; Score 34; DB 1; Length 936;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 RRRPRP 6
 |||||
 IDB 54 RRRPRP 59

RESULT 43
 YDOB_SCHPO STANDARD; PRT; 2052 AA.
 AC 013731;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 GN HYPOTHEICAL_234.0 KDA PROTEIN C15A10.11 IN CHROMOSOME I.
 GN SPAC15A10.11.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: SOME TO YEAST UBR1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z97208; CAB10108.1; -
 DR Hypothetical protein; Transmembrane.
 KW TRANSMEM 1024 1044 POTENTIAL.
 FT SEQUENCE 2052 AA; 234041 MW; 2AFA9E9E991D0453 CRC64;

Query Match
 Best Local Similarity 70.8%; Score 34; DB 1; Length 2052;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 6
 |||||
 IDB 554 RRRPRP 559

RESULT 44
 ORP1_HUMAN STANDARD; PRT; 2156 AA.
 ID ORP1_HUMAN
 AC P56715;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE OXYGEN-REGULATED PROTEIN 1 (RETINITIS PIGMENTOSA RPL PROTEIN)
 DE (RETINITIS PIGMENTOSA 1 PROTEIN).
 GN RPL OR ORP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS H-872; Y-985; T-1670; P-1691; Y-2033.
 RC TISSUE=Retina;
 RX MEDLINE=99318096; PubMed=10391212;
 RA Sullivan L.S., Heckenlied J.R., Boyne S.J., Zuo J., Hilde M.A.,
 Gal A., Denton M., Ingelhearn C.F., Blanton S.H., Dalgner S.P.;
 RT "Mutations in a novel retina-specific gene cause autosomal dominant
 RT retinitis pigmentosa.";
 RL Nat. Genet. 22:255-259(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=99318095; PubMed=10391211;
 RA Pierce E.A., Quinn T., Meenan T., McGee T.L., Berson E.L., Dryja T.P.;
 RT "Mutations in a gene encoding a new oxygen-regulated photoreceptor
 RT protein cause dominant retinitis pigmentosa.";
 RL Nat. Genet. 22:248-254(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99330563; PubMed=10401003;
 RA Guillonneau X., Piriev N.I., Danciger M., Kozak C.A., Cideciyan A.V.,
 RA Jacobson S.G., Farber D.B.;
 RT "A nonsense mutation in a novel gene is associated with retinitis
 RT pigmentosa in a family linked to the RPL locus.";
 RL Hum. Mol. Genet. 8:1541-1546(1999).
 CC -1- FUNCTION: COULD HAVE A ROLE IN THE DIFFERENTIATION OF
 CC PHOTORECEPTOR CELLS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN RETINA. NOT EXPRESSED IN HEART,
 CC BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, SPLEEN AND
 CC PANCREAS.
 CC -1- DISEASE: DEFECTS IN RPL CAUSE RETINITIS PIGMENTOSA FORM 1 (RPL); A
 CC DISEASE CHARACTERIZED BY CONSTRUCTION OF THE VISUAL FIELDS, NIGHT
 CC BLINDNESS, AND FUNDUS CHANGES. THE DISEASE SEEMS TO BE ASSOCIATED
 CC WITH TRUNCATED (STOP OR FRAMESHIFT MUTATIONS) FORMS OF THE
 CC PROTEIN.
 CC -1- SIMILARITY: CONTAINS 2 DOUBLECORTIN DOMAINS.
 CC -1- DATABASE: NAME=RetNet;
 CC NOTE=Retinal information network;
 CC WWW="http://www.sph.uth.tmc.edu/retnet/".
 CC -1- DATABASE: NAME=Mutations of the RPL gene;
 CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.itpa.org/sci-news/rplmt.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF143226; AAD44197.1; -
 DR EMBL: AF143224; AAD44197.1; JOINED.
 DR EMBL: AF143225; AAD44197.1; JOINED.
 DR EMBL: AF143222; AAD44198.1; -
 DR EMBL: AF141021; AAD42072.1; -
 DR EMBL: AF152242; AAD46774.1; -
 DR EMBL: AF152240; AAD46774.1; JOINED.

DR EMBL: AF152241; AAD46774.1; JOINED.
DR MIM: 603937; -
DR PROSITE: PS03030; DC: 2.
KW Vision; Retinitis pigmentosa; Retinal protein; Polymorphism; Repeat.
FT DOMAIN 36 118 DOUBLECORTIN.
FT DOMAIN 154 233 DOUBLECORTIN.
FT DOMAIN 268 273 POLY-SER.
FT DOMAIN 671 675 POLY-LYS.
FT DOMAIN 1687 1691 POLY-SER.
FT VARIANT 872 872 R -> H.
FT VARIANT 985 985 /FTid-VAR_007810.
FT VARIANT 1670 1670 N -> Y.
FT VARIANT 1691 1691 /FTid-VAR_007811.
FT VARIANT 1691 1691 A -> T.
FT VARIANT 1691 1691 /FTid-VAR_007812.
FT VARIANT 2033 2033 S -> P.
FT VARIANT 2033 2033 /FTid-VAR_007813.
FT VARIANT 2033 2033 C -> Y.
SQ SEQUENCE 2156 AA; 240659 MW; 55AEDBEC43D6A507 CRC64;
/FTid-VAR_007814.

Query Match
Best Local Similarity 70.8%; Score 34; DB 1; Length 2156;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 5.5e+02;

OY 1 RRRPRP 6
Db 121 RRRPRP 126

RESULT 45

E311_ADE02 STANDARD; PRT; 101 AA.
AC P24935;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE EARLY E3A 11.6 KDA GLYCOPROTEIN.
OS Human adenovirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=81053687; PubMed=6253880;
RA Herisse J, Coutouis G, Galibert F;
RL "Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome.";
Nucleic Acids Res. 8:2173-2192(1980).
RN [2]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=93079877; PubMed=1448922;
RA Scaria A, Tollefson A.E., Saha S.K., Wold W.S.M.;
RL "The E3-11.6K protein of adenovirus is an Asn-glycosylated integral
membrane protein that localizes to the nuclear membrane.";
Virology 191:743-753(1992).
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
KDA, AD5 E3A 10.5 KDA, AD3 E3 9 KDA, AND AD7 E3 7.7 KDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01917; AAA92222.1; -
DR Early protein; Glycoprotein; Transmembrane.
FT TRANSMEM 41 62 POTENTIAL.
FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 101 AA; 11644 MW; FB89FCC6E921E84B CRC64;

Query Match
Best Local Similarity 68.8%; Score 33; DB 1; Length 101;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPRP 7
Db 64 RRRPRP 70

Search completed: September 24, 2001, 10:09:07
Job time: 227 sec

ID 09FW62 PRELIMINARY; PRT; 136 AA.
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL 14.4 KDA PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Mofatt K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNB0094K03 genomic sequence."
 RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC069145; AAC16854.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 136 AA; 14435 MW; D3ACC33BBB362C91 CRC64;

Query Match 85.4%; Score 41; DB 10; Length 136;
 Best Local Similarity 100.0%; Pred. No. 9.3; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

OY 1 RRRPRP 7
 |||||
 DB 99 RRRPRP 105

RESULT 3
 ID 09SM77 PRELIMINARY; PRT; 168 AA.
 AC 09SM77;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ZH00008.1.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. INDICA;
 RA Hong G., Zhou B.;
 RT "Oryza sativa genomic DNA, chromosome 4, clone: t17804."
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL117265; CAB5416.1; -.
 SQ SEQUENCE 168 AA; 18668 MW; B9A19B483840F1A3 CRC64;

Query Match 85.4%; Score 41; DB 10; Length 168;
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

OY 1 RRRPRP 7
 |||||
 DB 56 RRRPRP 62

RESULT 4
 ID 042232 PRELIMINARY; PRT; 278 AA.
 AC 042232;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE Y-BOX BINDING PROTEIN A (FRAGMENT).
 GN YB-A.
 OS Columba livia (Domestic Pigeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
 OX NCBI_TaxID=8932;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Pratt S.L., Horseman N.D.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 DR EMBL: AF022769; AAB80761.1; -.
 DR HSSP: P41016; 1C90.
 DR InterPro: IPR002059; -.
 DR Pfam: PF00313; CSD; 1.
 DR PRINTS: PR00050; COLDSHOCK.
 DR SMART: SM00357; CSP; 1.
 KW DNA-binding; Transcription regulation.
 FT NON_TER 1
 SQ SEQUENCE 278 AA; 29614 MW; B51D26D012DF9410 CRC64;

Query Match 85.4%; Score 41; DB 13; Length 278;
 Best Local Similarity 100.0%; Pred. No. 18; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

OY 1 RRRPRP 7
 |||||
 DB 242 RRRPRP 248

RESULT 5
 ID 09LW27 PRELIMINARY; PRT; 300 AA.
 AC 09LW27;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE SIMILAR TO ORYZA SATIVA ROOT-SPECIFIC RCC3 MRNA.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
 clone: p0644B06."
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AP001129; BAA90617.1; -.
 DR InterPro: IPR000528; -.
 DR InterPro: IPR003612; -.
 DR Pfam: PF00279; LRP; 1.
 DR SMART: SM00499; AAI; 1.
 SQ SEQUENCE 300 AA; 30134 MW; B3A7C2E42CA3425A CRC64;

Query Match 85.4%; Score 41; DB 10; Length 300;
 Best Local Similarity 100.0%; Pred. No. 19; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

OY 1 RRRPRP 7
 |||||
 DB 178 RRRPRP 184

RESULT 6
 ID 09N103 PRELIMINARY; PRT; 304 AA.
 AC 09N103;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE Y-BOX PROTEIN ZONAB-A.
 GN ZONAB.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20253078; PubMed-10790369;
 RA Balda M.S., Matter K.;
 RT "The tight junction protein ZO-1 and an interacting transcription
 factor regulate Ehb-2 expression.";
 RL EMBO J. 19:2024-2033(2000).
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 DR EMBL; AF171061; AAF72335.1; -;
 DR InterPro; IPR002059; -;
 DR Pfam; PF00313; CSD; 1.
 DR PRINTS; PR00050; COLDSHOCK.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR SMART; SM00357; CSP; 1.
 KW DNA-binding; Transcription regulation.
 DR SMART; SM00357; CSP; 1.
 KW DNA-binding; Transcription regulation.
 SQ SEQUENCE 304 AA; 31939 MW; C49E737720F5202D CRC64;

Query Match 85.4%; Score 41; DB 6; Length 304;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
 DB 268 RRRPRP 274
 |||||

RESULT 7
 ID 09XCG4 PRELIMINARY; PRT; 361 AA.
 AC 09XCG4;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, last annotation update)
 DE HYPOTHEORETICAL 40.2 KDA PROTEIN.
 OS Mycobacterium avium.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1764;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2151;
 RC Eckstein T.M., Lambert M.L., Brennan P.J., Belisle J.T., Inamine J.M.;
 RT "Identification of a gene cluster involved in glycopeptidolipid
 biosynthesis and of a gene cluster encoding daunorubicin resistance in
 two strains of Mycobacterium avium serovar 2.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF143772; AAD44199.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 361 AA; 40208 MW; AD01DBE825C1C9EA CRC64;

Query Match 85.4%; Score 41; DB 2; Length 361;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
 DB 32 RRRPRP 38
 |||||

DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE Y-BOX PROTEIN ZONAB-B.
 GN ZONAB.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20253078; PubMed-10790369;
 RA Balda M.S., Matter K.;
 RT "The tight junction protein ZO-1 and an interacting transcription
 factor regulate Ehb-2 expression.";
 RL EMBO J. 19:2024-2033(2000).
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
 DR EMBL; AF171062; AAF72336.1; -;
 DR InterPro; IPR002059; -;
 DR Pfam; PF00313; CSD; 1.
 DR PRINTS; PR00050; COLDSHOCK.
 DR PRODOM; PD000621; -; 1.
 DR PROSITE; PS00352; COLD_SHOCK; 1.
 DR SMART; SM00357; CSP; 1.
 KW DNA-binding; Transcription regulation.
 SQ SEQUENCE 372 AA; 39879 MW; 33E579077CF2B969 CRC64;

Query Match 85.4%; Score 41; DB 6; Length 372;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
 DB 336 RRRPRP 342
 |||||

RESULT 9
 ID 066852 PRELIMINARY; PRT; 602 AA.
 AC 066852;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE PRE-TERMINAL PROTEIN.
 OS fowl adenovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Adenovirus.
 OX NCBI_TaxID=31540;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEKOTYPE 10;
 RA McCoy R.J., Sheppard M.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Sheppard M., Werner W., McCoy R., Johnson M.A.;
 RL Arch. Virol. 0:0-0(1997).
 DR EMBL; U26220; AAB02174.1; -;
 DR EMBL; AF007577; AAB8665.1; -;
 DR InterPro; IPR003391; -;
 DR Pfam; PF02459; Adeno_terminal; 1.
 SQ SEQUENCE 602 AA; 70392 MW; EB4827B4390894F CRC64;

Query Match 85.4%; Score 41; DB 14; Length 602;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
 DB 338 RRRPRP 344
 |||||

RESULT 8
 ID 09N102 PRELIMINARY; PRT; 372 AA.
 AC 09N102;
 DB 09N102;

```

RESULT 10
ID 010436 PRELIMINARY; PRT: 652 AA.
AC 010436;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE DNA POLYMERASE (E2B DNA POL) AND PTP (E2B PTP) GENES, PARTIAL CDS, AND
DE 52,55K (U1 52,55K), P11A (U1 P11A), III (U2 III), P11I (U2 P11I),
DE PUTATIVE 30 KDA PROTEIN (U2 COMP. ORF), V (U2 V) AND PX (U2 PX) GENES,
DE COMPLETE CDS (FRAGMENT).
DE E2B PTP.
GN E2B PTP.
OS Mouse adenovirus type 1 (MAV-1).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10530;
RN [1]
RP SEQUENCE FROM N.A.
RA Meisner J.D., Hirsch G.N., Larue E.A., Fulcher R.A., Spindler K.R.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: U95843; AAB53751.1; -.
DR InterPro: IPR003391; -.
DR Pfam: PF02459; Adeno_Terminal; 1.
FT NON_TER 1
SQ SEQUENCE 652 AA; 74663 MW; 1400CF2D2E266D0B CRC64;

Query Match 85.4%; Score 41; DB 14; Length 652;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRPP 7
   |||||
Db 387 RRRPP 393

RESULT 11
ID 070220 PRELIMINARY; PRT: 400 AA.
AC 070220;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE FORK HEAD TRANSCRIPTION FACTOR.
GN HFHL OR HFH-1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129; TISSUE=KIDNEY;
RX MEDLINE=98392851; PubMed=9726250;
RA Frank S., Zoll B.;
RT "Mouse HNF-3/fork head homolog-1-like gene: structure, chromosomal
RT location, and expression in adult and embryonic kidney.";
RL DNA Cell Biol. 17:679-688(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=129; TISSUE=KIDNEY;
RX Frank S., Zoll B.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=129; TISSUE=KIDNEY;
RC Pasche B., Bieller A., Zoll B.;
RA Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF010405; AAC12973.2; -.
DR HSSP: O63245; 2HFH.
DR MGD: MGI:1298328; HFHL.
DR InterPro: IPR001766; -.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR PROSITE: PS00657; FORKHEAD.
DR SMART: SM00339; FORK_HEAD_1; 1.
DR PROSITE: PS00657; FORK_HEAD_2; 1.
DR SMART: SM00339; FORK_HEAD_3; 1.
SQ SEQUENCE 400 AA; 41444 MW; 99318026C8E2AD19 CRC64;

Query Match 83.3%; Score 40; DB 11; Length 400;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 8
   |||||
Db 112 RRRPP 118

RESULT 13
ID 09J18 PRELIMINARY; PRT: 400 AA.
AC 09J18;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HFHL.
GN HFHL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129;
RC Hong H.-K., Noveroske J.K., Justice M.J., Chakravarti A.;
RA "The winged helix/forkhead transcription factor Hfh1 gene regulates
RT hair keratinization in satlin mice.";

```

```

DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
DR SMART: SM00339; FH; 1.
SQ SEQUENCE 400 AA; 41342 MW; E6C3B5F53FC42B CRC64;

Query Match 83.3%; Score 40; DB 11; Length 400;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 8
   |||||
Db 112 RRRPP 118

RESULT 12
ID 09J17 PRELIMINARY; PRT: 400 AA.
AC 09J17;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE WINGED HELIX/FORKHEAD TRANSCRIPTION FACTOR HFHL.
GN HFHL.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong H.-K., Noveroske J.K., Justice M.J., Chakravarti A.;
RT "The winged helix/forkhead transcription factor Hfh1 gene regulates
RT hair keratinization in satlin mice.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF153193; AAF73234.1; -.
DR InterPro: IPR001766; -.
DR Pfam: PF00250; Fork_head; 1.
DR PRINTS: PR00053; FORKHEAD.
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
DR PROSITE: PS50039; FORK_HEAD_3; 1.
DR SMART: SM00339; FH; 1.
SQ SEQUENCE 400 AA; 41444 MW; 99318026C8E2AD19 CRC64;

Query Match 83.3%; Score 40; DB 11; Length 400;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 8
   |||||
Db 112 RRRPP 118

RESULT 13
ID 09J18 PRELIMINARY; PRT: 400 AA.
AC 09J18;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HFHL.
GN HFHL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129;
RC Hong H.-K., Noveroske J.K., Justice M.J., Chakravarti A.;
RA "The winged helix/forkhead transcription factor Hfh1 gene regulates
RT hair keratinization in satlin mice.";

```


3L Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF154426; AAF74524.1; -
 DR InterPro: IPR001766; -
 DR Pfam: PF00250; Fork_head; 1.
 DR PRINTS: PR00053; FORKHEAD.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; 1.
 DR PROSITE: PS50039; FORK_HEAD_3; 1.
 DR SMART: SM00339; FH; 1.
 SQ SEQUENCE 400 AA; 41368 MW; F286AEFF81D4759B CRC64;

Query Match
 Best Local Similarity 83.3%; Score 40; DB 11; Length 400;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPPPY 8
 DB 112 RRRPPPY 118

RESULT 14
 O9NS06 PRELIMINARY; PRT; 402 AA.
 AC O9NS06;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE WINED HELIX/FORKHEAD TRANSCRIPTION FACTOR.
 GN HFI1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Hong H.-K., Noveroske J.K., Justice M.J., Chakravarti A.;
 RT "The winged helix/forkhead transcription factor Hfi1 gene regulates
 hair keratinization in satin mice."
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF153341; AAF75586.1; -
 DR InterPro: IPR001766; -
 DR Pfam: PF00250; Fork_head; 1.
 DR PRINTS: PR00053; FORKHEAD.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; 1.
 DR PROSITE: PS50039; FORK_HEAD_3; 1.
 DR SMART: SM00339; FH; 1.
 SQ SEQUENCE 402 AA; 41485 MW; FD2EEFF99D8848E77 CRC64;

Query Match
 Best Local Similarity 83.3%; Score 40; DB 4; Length 402;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPPPY 8
 DB 115 RRRPPPY 121

RESULT 15
 O9RE11 PRELIMINARY; PRT; 89 AA.
 AC O9RE11;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 10.6 KDA PROTEIN.
 OS Lactobacillus casei.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1582;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 393;
 RA Dossoumet V., Deutscher J.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y18948; CAB65149.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 89 AA; 10573 MW; E7ED1D5E7BFD668 CRC64;

Query Match
 Best Local Similarity 81.2%; Score 39; DB 2; Length 89;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPPY 8
 DB 56 RRRPPPY 63

RESULT 16
 O9TY3 PRELIMINARY; PRT; 488 AA.
 AC O9TY3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE K08B4.5 PROTEIN.
 GN K08B4.5
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favallo A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Rohlfing T., Murray J., Antonio B.;
 RT "The sequence of C. elegans cosmid K08B4."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF100663; AAC68982.1; -
 DR InterPro: IPR001394; -
 DR Pfam: PF00443; UCH-2; 1.
 DR PROSITE: PS50235; UCH_2_3; 1.
 SQ SEQUENCE 488 AA; 56286 MW; 33828EAF9F0D082B CRC64;

Query Match
 Best Local Similarity 81.2%; Score 39; DB 5; Length 488;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPPY 8
 DB 345 RRRPPPY 352

```

RESULT 17
ID Q9UF45 PRELIMINARY; PRT; 237 AA.
AC Q9UF45;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DE HYPOTHEICAL 27.1 KDA PROTEIN (FRAGMENT).
GN DKF2P434F178.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Bloeker H., Boecker M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL33615, CAB63743.1; -.
RN NCBI_TaxID=9606;
PT NON-TER
SQ SEQUENCE 237 AA; 27110 MW; F3753DBD81E8C368 CRC64;

Query Match
Best Local Similarity 79.2%; Score 38; DB 4; Length 237;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
Db 224 RRRPKRP 230

RESULT 18
ID Q9ZBR7 PRELIMINARY; PRT; 322 AA.
AC Q9ZBR7;
DT 01-MAY-1999 (TRENBLREL. 10, Created)
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
DE PUTATIVE THIAMINE MONOPHOSPHATE KINASE.
GN SC7A1.06.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Murphy L., Harris D.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RL "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome."
DR EMBL: AL034447; CAA22405.1; -.
DR InterPro: IPR000728; -.
DR Pfam: PF00586; ATRS; 1.
KW kinase.
SQ SEQUENCE 322 AA; 33610 MW; DE538BF5BE0F87B6 CRC64;

```

```

Query Match
Best Local Similarity 79.2%; Score 38; DB 2; Length 322;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPRP 8
Db 192 RRRPPY 198

RESULT 19
ID Q9P2B7 PRELIMINARY; PRT; 527 AA.
AC Q9P2B7;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DE KIAA1430 PROTEIN (FRAGMENT).
GN KIAA1430.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RL "Prediction of the coding sequences of unidentified human genes. XVI.
for large proteins in vitro."
DR EMBL: AB037851; BAA92668.1; -.
RN NCBI_TaxID=9606;
PT NON-TER
SQ SEQUENCE 527 AA; 58896 MW; 082CF438BB9F4F0A CRC64;

Query Match
Best Local Similarity 79.2%; Score 38; DB 4; Length 527;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRPRP 7
Db 514 RRRPKRP 520

RESULT 20
ID P74024 PRELIMINARY; PRT; 533 AA.
AC P74024; O08043;
DT 01-FEB-1997 (TRENBLREL. 02, Created)
DT 01-FEB-1997 (TRENBLREL. 02, Last sequence update)
DE 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
DE HYDROGENASE SUBUNIT.
GN HOX OR SL121.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RL "Sequence analysis of the genome of the unicellular cyanobacterium
Synecocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions."
RN [3]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-97135082; PubMed-8980640;
RA Appel J., Schulz R.;
RT "Sequence analysis of an operon of a NAD(P)-reducing nickel
hydrogenase from the cyanobacterium Synechocystis sp. PCC 6803 gives
additional evidence for direct coupling of the enzyme to NAD(P)H-
dehydrogenase (complex I).";
RL Biochim. Biophys. Acta 1298:141-147(1996).
DR EMBL: D90911; BAA18097.1; -.
DR EMBL: X97610; CAA66209.1; -.
DR InterPro: IPR001949; -.
DR Pfam: PF01512; ComplexI_51k; 1.
DR PROSITE: PS00645; COMPLEXI_51K_2; 1.
SQ SEQUENCE 533 AA; 57774 MW; 2DC4366CFC8652A CRC64;

OY 3 RRRPPY 8
Db 307 RRRPPY 312

Query Match 79.2%; Score 38; DB 2; Length 533;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21
Db 307 RRRPPY 312

RESULT 21
O9SUT1 PRELIMINARY; PRT; 700 AA.
AC O9SUT1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE PUTATIVE UBIQUITIN ACTIVATING ENZYME.
GN AT2G31470.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicaceae; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

FN [1]
FP SEQUENCE FROM N.A.
FC STRAIN-CV. COLOMBIA;
RA MEDLINE-20083487; PubMed-10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-T., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.-J., Rinning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Umeyam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL: AC006841; AAD23691.1; -.
DR InterPro: IPR00127; -.
DR InterPro: IPR00205; -.
DR InterPro: IPR00594; -.
DR Pfam: PF00889; Thir_family; 1.
DR Pfam: PF02134; UBACT_repeat; 1.
SQ SEQUENCE 700 AA; 78525 MW; B96A697287F1C1B CRC64;

OY 1 RRRPPY 8
Db 257 RRRPPY 264

Query Match 79.2%; Score 38; DB 10; Length 700;
Best Local Similarity 87.3%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 22
Db 257 RRRPPY 264

RESULT 22
O9PFK8 PRELIMINARY; PRT; 224 AA.
AC O9PFK8;
O9PFK8;
O9PFK8;

```

```

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN XF0649.
GN XF0649.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9A5C;
RX MEDLINE-20365717; PubMed-10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohenisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AE003909; AAF83459.1; -.
DR InterPro: IPR002610; -.
DR Pfam: PF01694; Rhomboid; 1.
KW Hypothetical protein.
SQ SEQUENCE 224 AA; 24904 MW; 5DA58727CCA0DD4 CRC64;

OY 2 RRRPPY 8
Db 213 RRRPPY 219

Query Match 77.1%; Score 37; DB 2; Length 224;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 23
Db 213 RRRPPY 219

RESULT 23
O9HXC9 PRELIMINARY; PRT; 276 AA.
AC O9HXC9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROBABLE SHORT-CHAIN DEHYDROGENASE.
GN PA3883.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PA01;

```

RA MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goitry L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004805; AGO07270.1; -
 DR InterPro: IPR002198; -
 DR InterPro: IPR002347; -
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00081; GDRHD.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 SO SEQUENCE 276 AA; 29427 MW; CA7458A8C6CA15A3 CRC64;

Query Match 77.1%; Score 37; DB 2; Length 276;
 Best Local Similarity 85.7%; Pred. No. 78;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPP 7
 ID :|||||
 DB 233 ORRRPP 239

RESULT 24
 O9LHG8 PRELIMINARY; PRT; 398 AA.

AC O9LHG8; 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE MOUSE AND HUMAN TUMOR SUSCEPTIBILITY GENE-LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RA Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX PubMed=10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety pl.
 RT TAC and BAC clones."
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP002047; BAB03147.1; -
 SO SEQUENCE 398 AA; 44715 MW; 659996F9784FDC20 CRC64;

Query Match 77.1%; Score 37; DB 10; Length 398;
 Best Local Similarity 85.7%; Pred. No. 11e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRRPP 7
 ID :|||||
 DB 161 RRRRPP 167

RESULT 25
 O9MAV4 PRELIMINARY; PRT; 70 AA.
 ID O9MAV4;
 AC O9MAV4;

DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE F2401.6.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F2401 from chromosome
 RT I."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thayerl A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC003113; AAF70835.1; -
 DR InterPro: IPR00216; -
 DR InterPro: IPR002965; -
 DR PRINTS: PR01217; PRICHEXTENS.
 DR PRINTS: PR00239; RHODOPENTAIL.
 SO SEQUENCE 70 AA; 7633 MW; 05AA04E5484D4D48 CRC64;

Query Match 75.0%; Score 36; DB 10; Length 70;
 Best Local Similarity 85.7%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRRPP 8
 ID :|||||
 DB 64 RRRRPP 70

RESULT 26
 O9VB37 PRELIMINARY; PRT; 108 AA.
 ID O9VB37;
 AC O9VB37;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE CG12510 PROTEIN.
 GN CG12510.

OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brostein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gidbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003760; AAF56706.1; -
 DR FlyBase: FBgn0039515; CG12510.
 SQ SEQUENCE 108 AA; 10799 MW; 29992588920FD967 CRC64;

Query Match 75.0%; Score 36; DB 5; Length 108;
 Best Local Similarity 85.7%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPPPP 7
 I I I I I
 Db 92 RHRPPPP 98

RESULT 27
 AC 024156 PRELIMINARY; PRT; 129 AA.
 AC 024156;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE GLYCINE-RICH PROTEIN (FRAGMENT).
 DE NTGRP15.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;

OC Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fuertemberg S.T., Buccigaglia P.A., Smith A.G.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U74325; AAB18261.1; -
 DR Mendel: 25642; Nicita;343;25642.
 FT NON_TER 1
 SQ SEQUENCE 129 AA; 12251 MW; C6F5A54DE29F7597 CRC64;

Query Match 75.0%; Score 36; DB 10; Length 129;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPPPP 8
 I I I I I
 Db 28 RRRPPPP 34

RESULT 28
 AC 094184 PRELIMINARY; PRT; 171 AA.
 AC 094184;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
 DE COSMID F16F9.
 GN F16F9.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peldotinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifen L., Roopra A., Saunders D., Showkseen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Fulton B.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U67956; AAB07690.1; -
 DR EMBL: U67956; AAB07690.1; -
 SQ SEQUENCE 171 AA; 19114 MW; ECF084C0B3D4ED9B CRC64;

Query Match 75.0%; Score 36; DB 5; Length 171;
 Best Local Similarity 85.7%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPPPP 8
 I I I I I
 Db 12 RRRPPPP 18

RESULT 29
 ID 054364 PRELIMINARY; PRT; 190 AA.
 AC 054364;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE (78-11).
 RN LMBK.
 GN Streptomyces lincolnensis.
 OS Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1915;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=78-11;
 RC MEDLINE=96020646; PubMed=8577249;
 RA Peschke U., Schmidt H., Zhang H.Z., Piepersberg W.;
 RT "Molecular characterization of the lincomycin-production gene cluster
 of Streptomyces lincolnensis 78-11.";
 RL MOL. Microbiol. 16:1137-1156(1995).
 DR EMBL; X79146; CA55757.1; -;
 DR InterPro; IPR001454; -;
 DR Pfam; PF00702; Hydrolase; 1.
 SO SEQUENCE 190 AA; 20298 MW; 7EF26984BED18DE3 CRC64;

Query Match 75.0%; Score 36; DB 2; Length 190;
 Best Local Similarity 85.7%; Pred. No. 81;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPPY 8
 ||| |||
 DB 164 RRRPPY 170

RESULT 30
 ID 039716 PRELIMINARY; PRT; 226 AA.
 AC 039716;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE PRECURSOR PROTEIN OF EEP-2.
 OS Eisenia foetida (Common branding worm) (Common dung-worm).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 OC Lumbricidae; Lumbricidae; Eisenia.
 OX NCBI_TaxID=6396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawano T., Oumi T., Ukena K., Ikeda T., Matsushima O., Minakata H.;
 RT "Gene structure of the GNG-peptides derived from Annelida and the
 gene products.";
 RL (In) Nishi N. (eds.);
 RL Peptide chemistry, pp.321-324, Protein Research Foundation,
 Osaka (1995).
 RL EMBL; D50862; BAA09458.1; -;
 DR Mendel; 11730; Eise;1032;11730.
 SO SEQUENCE 226 AA; 25588 MW; 7997C50309743C5B CRC64;

Query Match 75.0%; Score 36; DB 5; Length 226;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 ||| |||
 DB 118 RRRPP 123

RESULT 31
 O9RBX8

ID O9RBX8 PRELIMINARY; PRT; 228 AA.
 AC O9RBX8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 24.1 KDA PROTEIN.
 OS Pseudomonas indigofera.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Vogesella.
 OX NCBI_TaxID=45465;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC19706;
 RA van de Loo F.J., Keese P., Llewellyn D.;
 RT "Structural and regulatory genes controlling indolizidine production in
 Vogesella indigofera: involvement of a peptide synthetase homolog.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
 DR EMBL; AF088856; AAD54003.1; -;
 DR InterPro; IPR001647; -;
 DR Pfam; PF00440; tetr; 1.
 DR PRINTS; PR00455; HPTETR.
 KW DNA-binding; Hypothetical protein; Transcription regulation.
 SO SEQUENCE 228 AA; 24140 MW; A928DE1AF404869B CRC64;

Query Match 75.0%; Score 36; DB 2; Length 228;
 Best Local Similarity 75.0%; Pred. No. 96;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPPY 8
 ||| |||
 DB 23 RRRPPY 30

RESULT 32
 ID 012396 PRELIMINARY; PRT; 248 AA.
 AC 012396;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE NEUROVIRULENCE FACTOR.
 OS human herpesvirus 1.
 GN Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10298;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RA McGeoch D.J.;
 RL Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RC MEDLINE=89036163; PubMed=2846760;
 RA Perry D.J., McGeoch D.J.;
 RT "The DNA sequences of the long repeat region and adjoining parts of
 the long unique region in the genome of herpes simplex virus type 1.";
 RL J. Gen. Virol. 69:2831-2846(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RC MEDLINE=86148504; PubMed=3005980;
 RA McGeoch D.J., Dolan A., Donald S., Brauer D.H.;
 RT "Complete DNA sequence of the short repeat region in the genome of
 herpes simplex virus type 1.";
 RL Nucleic Acids Res. 14:1727-1745(1986).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RC MEDLINE=92341080; PubMed=1321882;
 RA Dolan A., Mckie E., Maclean A.R., McGeoch D.J.;

RT Status of the ICP34.5 gene in herpes simplex virus type 1 strain
RT 17.":
RL J. Gen. Virol. 73:971-973(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-17;
RA Dolan A.;
RU Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: X14112; CA332285.1; -;
EMBL: X14112; CA33248.1; -;
DR EMBL: X14112; CA33248.1; -;
SQ SEQUENCE 248 AA; 26195 MW; BF216520CB379478 CRC64;

Query Match 75.0%; Score 36; DB 14; Length 248;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 11 RRRPP 16

RESULT 33
ID 09DOX5 PRELIMINARY; PRT; 252 AA.
AC 09DOX5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GAMMA1 34.5.
OS human herpesvirus 1.
CC viruses: dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=10298;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1GC;
RA Ling J.-Y., Chen T.-M., Stroop W.G.;
RT "A gamma1 34.5 null mutant of herpes simplex virus type 1 strain +GC
is neurovirulent."
RU Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF251158; AAG39340.1; -;
SQ SEQUENCE 252 AA; 26631 MW; 908AE67C49A52C56 CRC64;

Query Match 75.0%; Score 36; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 12 RRRPP 17

RESULT 34
ID 099949 PRELIMINARY; PRT; 280 AA.
AC 099949;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE 179115.1 PROTEIN (FRAGMENT).
GN KIOTOH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurtry A., Odell C.;
RU Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z92540; CAB06876.1; -;
HSP: P26205; ICBG.

DR InterPro: IPR001360; -;
DR Pfam: PF00232; Glyco_hydro_1; 2.
DR PROSITE: PS00653; GLYCOSYL_HYDROL_F1_2; UNKNOWN_1.
FT NON_TER 280 280
SQ SEQUENCE 280 AA; 30818 MW; 96A336F65482DCAC CRC64;

Query Match 75.0%; Score 36; DB 4; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 8 RRRPP 13

RESULT 35
ID 09UHR3 PRELIMINARY; PRT; 366 AA.
AC 09UHR3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE NASOPHARYNGEAL CARCINOMA SUSCEPTIBILITY PROTEIN LZ16.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA zhan F.H., Li G.Y.;
RT "Cloning and identification of LZ16 Gene in NPC."
RU Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF121775; AAF24125.1; -;
DR HSP: P55273; 1BD8.
DR InterPro: IPR002110; -;
DR Pfam: PF00023; ank; 4.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR SMART: SM00248; ANK; 1.
SQ SEQUENCE 366 AA; 40103 MW; 246D0D16C59FDB96 CRC64;

Query Match 75.0%; Score 36; DB 4; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 337 RRRPP 342

RESULT 36
ID 057567 PRELIMINARY; PRT; 406 AA.
AC 057567;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HEDGEHOG SEGMENT POLARITY HOMOLOG.
OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
OC Notoththalmus.
OX NCBI_TaxID=8316;
RN [1]
RP SEQUENCE FROM N.A.
RA Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;
RU Dev. Dyn. 0:0-0(1998).
CC -I- SIMILARITY: TO PURINE/PYRIMIDINE PHOSPHORIBOSYL TRANSFERASE
CC FAMILY.
DR EMBL: AF047466; AAC03108.1; -;
HSP: Q62226; LVHH.

```

DR InterPro: IPR000320; -
DR InterPro: IPR001657; -
DR InterPro: IPR001767; -
DR InterPro: IPR002375; -
DR InterPro: IPR003586; -
DR Pfam: PF01079; Hint; 1.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR PRODOM: PD003042; -; 1.
DR PROSITE: PS00103; PUR_Pyr_PR_TRANSFER; 1.
DR SMART: SM00305; HincC; 1.
DR GlycoSylTransferase: Transferase.
KW GlycoSylTransferase; 406 AA; 45072 MW; 5842CCADS314D4D0 CRC64;

SQ
Query Match
Best Local Similarity 75.0%; Score 36; DB 13; Length 406;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 33 RRRPP 38

RESULT 37
Q98845 PRELIMINARY; PRT; 502 AA.
ID Q98845;
AC Q98845;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE MATRIBETIC PEPTIDE RECEPTOR TYPE-C (NPR-C).
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GILL;
RA MEDLINE=95172050; PubMed=7867625;
RA Takashima A., Katafuchi T., Shibasaki M., Hagiwara H.,
RA Takei Y., Hirose S.;
RA "Cloning, properties, site-directed mutagenesis analysis of the
RT subunit structure, tissue distribution and regulation of expression of
RT the type-C eel natriuretic peptide receptor.";
RL Eur. J. Biochem. 227:673-680(1995).
DR EMBL; D31928; BAA06711.1; -
DR InterPro: IPR001170; -
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00255; NATPEPTIDER.
DR PROSITE: PS00458; ANF_RECEPTORS; 1.
DR Transmembrane.
KW SEQUENCE 502 AA; 56366 MW; 24F8BC6D96B58923 CRC64;

SQ
Query Match
Best Local Similarity 75.0%; Score 36; DB 13; Length 502;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRPP 7
DB 422 RRRPP 428

RESULT 38
Q916G2 PRELIMINARY; PRT; 524 AA.
ID Q916G2;
AC Q916G2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

```

```

DE PUTATIVE ANTHOCYANIN 5-O-GLUCOSYLTRANSFERASE.
GN P0406H10.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
OC Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0406H10.1";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002524; BAB07962.1; -
DR InterPro: IPR002213; -
DR Pfam: PF00201; UDPGT; 1.
KW Transferase
KW SEQUENCE 524 AA; 56014 MW; 833A61A10805F7E4 CRC64;

SQ
Query Match
Best Local Similarity 75.0%; Score 36; DB 10; Length 524;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRRPP 8
DB 12 RRRPP 18

RESULT 39
Q53913 PRELIMINARY; PRT; 527 AA.
ID Q53913;
AC Q53913;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE BETA-KETOACYL SYNTHASE HOMOLOGUE (FRAGMENT).
OS Streptomyces coriolicus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=28043;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94291952; PubMed=8020754;
RA Kuzek K., Mordarski M., Goodfellow M.;
RA "Distribution of oxoacyl synthase homology sequences within
RT Streptomyces DNA.";
RL FEMS Microbiol. Lett. 118:317-325(1994).
DR EMBL; I20249; AAA21754.1; -
FT NON_TER 1
FT NON_TER 527
FT SEQUENCE 527 AA; 56429 MW; 6574224B6A522D12 CRC64;

SQ
Query Match
Best Local Similarity 75.0%; Score 36; DB 2; Length 527;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRPP 7
DB 327 RRRPP 332

RESULT 40
Q9Y4F0 PRELIMINARY; PRT; 549 AA.
ID Q9Y4F0;
AC Q9Y4F0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE KLOTTHO PROTEIN (KL).
GN KLOTTHO.

```


OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98125543; PubMed=9464267;
 RA Matsumura Y., Aizawa H., Shiraki-Iida T., Nagai R., Kuro-O M.,
 Nabeshima Y.;
 RT Identification of the human klotho gene and its two transcripts
 encoding membrane and secreted klotho protein.";
 RL Biochem. Biophys. Res. Commun. 242:626-630(1998).
 DR EMBL: AB009667; BAA24941.1; -
 DR EMBL: AB009666; BAA24941.1; JOINED.
 DR HSSP: P26205; 1C8C.
 DR InterPro: IPR001360; -
 DR Pfam: PF00232; Glyco_hydro_1; 3.
 DR PRINTS: PR00131; GLHYDRLASE.
 DR PROSITE: PS00653; GLYCOSYL_HYDROL_F1_2; UNKNOWN_1.
 SO SEQUENCE 549 AA; 62165 MW; 5A05A64DB71C499A CRC64;

Query Match 75.0%; Score 36; DB 4; Length 549;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 7
 ID 8 RRRPP 13

RESULT 41

C9JG58 PRELIMINARY; PRT; 665 AA.
 AC 09JG58;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE ORF1.
 CS TTV-like mini virus.
 CC Viruses; ssDNA viruses; Circoviridae.
 OX NCBI_TaxID=93678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIMV-CLC156;
 RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
 RT Full or near-full length nucleotide sequences of TTV variants (types
 SANBAN and YONBAN) and the TTV-like mini virus (TIMV).";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB038627; BAA93601.1; -
 SO SEQUENCE 665 AA; 79465 MW; 3CAB89342D88B244 CRC64;

Query Match 75.0%; Score 36; DB 14; Length 665;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRRPP 8
 ID 18 RRRPP 25

RESULT 42

074935 PRELIMINARY; PRT; 700 AA.
 AC 074935;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ACYL-COA OXIDASE 2 (EC 1.3.3.6).
 GN ACO2.
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W29;
 RA Nicaud J.M.J.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W29;
 RA Le Clainche A.;
 RL Thesis (1997), Institut National Agronomique Paris-Grignon.
 DR EMBL: AJ001300; CA04660.1; -
 DR InterPro: IPR001005; -
 DR InterPro: IPR001993; -
 DR InterPro: IPR002655; -
 DR Pfam: PF01756; ACOX; 1.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 KW Oxidoreductase.
 SO SEQUENCE 700 AA; 78690 MW; 434C97B8A7D71E3E CRC64;

Query Match 75.0%; Score 36; DB 3; Length 700;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRRPP 8
 ID 670 RRRPP 676

RESULT 43

075048 PRELIMINARY; PRT; 903 AA.
 AC 075048;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE KIAA0460 PROTEIN (FRAGMENT).
 GN KIAA0460.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=98116662; PubMed=9455484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
 Nomura N., Ohara O.;
 RT Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain.";
 RL DNA Res. 4:345-349(1997).
 DR EMBL: AB007929; BAA32305.1; -
 FT NON_TER 1
 SO SEQUENCE 903 AA; 95633 MW; CB80B7268808420C CRC64;

Query Match 75.0%; Score 36; DB 4; Length 903;
 Best Local Similarity 71.4%; Pred. No. 3.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRPP 8
 ID 879 RRRPP 885

RESULT 44

09VC93 PRELIMINARY; PRT; 982 AA.
 AC 09VC93;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE CG6413 PROTEIN.
 GN CG6413.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132.
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhattacharya D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner V., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003747; AAF56281.1; -;
 DR FlyBase: FBgn0039183; CG6413.
 DR InterPro: IPR001900; -;
 DR Pfam: PF00773; RMB: 1.
 DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
 DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
 SQ SEQUENCE 982 AA; 112182 MW; 9C55A04867611FF4 CRC64;

Query Match 75.0%; Score 36; DB 5; Length 982;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RRRPRP 7
 DB 696 RRRPRP 702
 RESULT 45
 ID Q9UEI9 PRELIMINARY; PRT; 1012 AA.
 AC Q9UEI9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE KLOTIO PROTEIN (KL).
 GN KLOTIO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98125543; PubMed=9464267;
 RA Matsumura Y., Atzawa H., Shitaki-Tida T., Nagai R., Kuro-o M.,
 RA Nabeshima Y.;
 RT "Identification of the human Klotio gene and its two transcripts
 RT encoding membrane and secreted Klotio protein.";
 RL Biochem. Biophys. Res. Commun. 242:626-630(1998).
 DR EMBL: AB009667; BAA24940.1; -;
 DR EMBL: AB009666; BAA24940.1; JOINED.
 DR HSSP: P26205; ICBG.
 DR InterPro: IPR001360; -;
 DR Pfam: PF00232; Glyco_hydro.1; 6.
 DR PRINTS: PR00131; GLHYDRLASE1.
 DR PROSITE: PS00653; GLYCOSYL_HYDROL_F1_2; UNKNOWN.1.
 DR PROSITE: PS00653; GLYCOSYL_HYDROL_F1_2; UNKNOWN.1.
 SQ SEQUENCE 1012 AA; 116211 MW; 7A22C9CA1FE02E84 CRC64;

Query Match 75.0%; Score 36; DB 4; Length 1012;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RRRPRP 7
 DB 8 RRRPRP 13

Search completed: September 24, 2001, 10:07:44
 Job time: 144 sec
